OBJECTS AND PERSONS

TRENTON MERRICKS

Objects and Persons

Objects and Persons

Trenton Merricks

CLARENDON PRESS · OXFORD
2001

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Trenton Merricks 2001
The moral rights of the authors have been asserted
Database right Oxford University Press (maker)
First published 2001
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Merricks, Trenton.
Objects and persons / Trenton Merricks.
p. cm.
Includes bibliographical references and index.
1. Ontology. 2. Whole and parts (Philosophy). 3. Agent (Philosophy).
4. Causation. I. Title.
BD396.M47 2001 110—dc21 00–068260
ISBN 0–19–924536–3

For Laura

Preface

Ontological discovery is not empirical. But ontologists do make discoveries. Empirical investigation might tell us that an object is perforated. But it won't discover that there are holes, surrounded by (or partially composing) perforated objects. Only a good argument could discover that. We can see that one thing is the same colour as another. But whether this implies the existence of a universal, present in both, must be resolved philosophically. Census-takers may count us. But only an ontologist can find out whether there is a number that numbers us.

Ontological discovery is not empirical. But ontologists do make discoveries. Or so say believers in ontology. And I believe. If seeing were believing, then by the end of this book you would believe too. For—assuming my arguments are successful—ontological discoveries follow.

I shall argue that there are no inanimate macrophysical objects such as statues or baseballs or rocks or stars. But my 'eliminativism' about these objects—like (so one might argue) controversial ontological claims about holes, universals, and numbers—is consistent with the empirically established facts. This is because, as we shall see, I agree with my opponents that there are microscopica arranged in various ways, such as 'statuewise', 'baseballwise', 'rockwise', and 'starwise'.

Chapter 1 explains eliminativism in detail and addresses challenges to its coherence. Chapter 2 raises a number of considerations that motivate eliminativism. In Chapter 3 we find that, if things like statues and baseballs existed, everything they allegedly cause would be caused by their parts; if statues

and baseballs existed, they would be—at best—wholly causally redundant. This, I argue, leads to their elimination.

So much for what there is not. The next three chapters deal with what there is, focusing primarily on us human organisms. Chapter 4 argues that we have 'non-redundant' causal powers—we can cause things not causally overdetermined by our proper parts—and that this keeps us from being eliminated by the arguments of Chapter 3. Chapter 5 shows why Chapter 2's considerations do not motivate eliminating us.

Chapter 4's arguments work only if some conscious mental states are causally efficacious. So Chapter 6 blocks a serious argument for mental epiphenomenalism. That chapter also argues, among other things, that incompatibilists about free will should endorse the claim that we have non-redundant causal powers of the sort defended in Chapter 4. At least, incompatibilists should endorse this if they believe that we are human organisms who act freely.

The seventh and final chapter argues that, though both believe falsely, someone who believes in statues (and baseballs and rocks and the rest) is better off than someone who believes in unicorns.

My ontology is, more or less, in the tradition of those (arguably Aristotle's, obviously van Inwagen's) that endorse organisms and eliminate inanimate composite objects. But only more or less. For while I deny the existence of inanimate macroscopica—statues, baseballs, rocks, stars, etc.—their problem is not that they are inanimate. Their problem is, among other things, that were they to exist, their causal powers would be at best redundant. (Thus it makes no real difference to my ontology if there happen to be some exotic inanimate macroscopica, just so long as they have non-redundant causal powers.) And I defend our existence—not on the grounds that we are alive—but on the grounds that, among other things, we have non-redundant causal powers. I am happy to eliminate any alleged organisms that, if they

existed, would at most cause only what their parts overdetermine.

I wrote most of this book during the academic year 1999/2000, while enjoying a fellowship from the National Endowment for the Humanities and a leave from Virginia Commonwealth University. I am grateful to the NEH and to VCU for their generosity.

Close ancestors of two short arguments in Chapter 1, and of an argument in Chapter 4, have already been published. They can be found in 'Composition as Identity, Mereological Essentialism, and Counterpart Theory', " 'No Statues' ", and 'Against the Doctrine of Microphysical Supervenience'. The first two articles appeared in the *Australasian Journal of Philosophy,* the third in *Mind*. I presented other arguments from this book at the University of Virginia (1999), the Eastern Division APA Symposium on Ontology (1999), and Notre Dame's awkwardly named but smoothly run Mighty Midwestern Metaphysical Mayhem IV (1999).

I received a great deal of help on this book. I here offer my thanks to the many who made suggestions, raised objections, and responded to my questions. Thanks also go to Randy Carter, Tamar Gendler, John Hawthorne, Mark Heller, Jaegwon Kim, Al Plantinga, Thomas Williams, and Dean Zimmerman. They gave careful attention to extensive stretches of one or another draft of this book. And I am especially grateful to the following who provided valuable comments on (in some cases multiple drafts of) the entire manuscript: Mike Bergmann, Jonathan Lowe, Gene Mills, Mark Murphy, Eric Olson, Mike Rea, Alan Sidelle, and Ted Sider.

T. M.

Richmond, Virginia

Contents

1. Explaining Eliminativism 1
   I. Eliminativism: The Basic Idea 2
   II. Eliminativism: Not as Bad as you might Think 8
   III. The Linguistic Charge of Contradiction 12
   IV. The Metaphysical Charge of Contradiction 20
   V. Conclusion 29
2. Considerations in Favour of Eliminativism 30
   I. The Water in the Pool 30
   II. The Sorites Game 32
   III. The Statue and the Lump 38
   IV. Brains and Thinkers 47
   V. Conclusion 53
3. Epiphenomenalism and Eliminativism 56
   I. The Causal Principle 57
   II. Atomic Causation 59
   III. Causal Overdetermination 66
   IV. The Moral of the Overdetermination Argument 79
   V. Conclusion 83
4. Surviving Eliminativism 85
   I. Step One 89
   II. Conscious Mental Properties and Premiss (1*a*) 93

III. Objections to the Defence of Premiss (1*a*) 96
IV. Step One Again 104
V. Step Two 107
VI. On What Composite Objects Exist 114
VII. Conclusion 116
5. Considerations in Favour of Eliminating Us? 118
I. Persons and the Water in the Pool 119
II. Persons and the Sorites Game 124
III. Statues, Lumps, and Persons 130
IV. Brains, Thinkers, and Persons 135
V. Conclusion 137
6. Mental Causation and Free Will 138
I. The Exclusion Argument(s) 138
II. Causal Overdetermination Again 146
III. The 'Bottom-Up' Threat to Free Will 155
IV. Conclusion 161
7. Belief and Practice 162
I. False Folk Beliefs 162
II. False Folk Beliefs are Nearly as Good as True: Justification 171
III. False Folk Beliefs are Nearly as Good as True: Practice 175
IV. And Yet I Often Say 'There are statues' 186
V. Conclusion 190
References 191
Index 201

1 Explaining Eliminativism

In this book I shall show that there are no books. Nor are there statues, rocks, tables, stars, or chairs. Indeed I shall argue that there are no inanimate macrophysical objects at all. Thus I shall argue against the existence of most of the objects alleged to exist by what we might call, to be trendy, 'folk ontology'. I shall, however, defend the existence of the folk themselves, and shall do so on the assumption that they are human organisms. As we shall see (Chapter 4, §VI), it will remain somewhat of an open question which other alleged organisms really exist.

I cannot fully explain why I deny the existence of, say, statues, but not human organisms, without arguing for the truth of my ontology. But the aim of this chapter is not to present those arguments; it is not to defend that ontology. That is the work of later chapters, starting with Chapter 2. For I cannot even begin to defend my ontology until I make clear what exactly that ontology *is*.

The claim that we human organisms exist is not likely to be misunderstood or assailed by objections that simply miss the mark. And so I won't dwell on that claim in this chapter. I will instead focus on my 'eliminativism' about non-living macrophysical objects. This sort of eliminativism is often, at least

initially, grossly misunderstood. This chapter aims to rectify the gross misunderstandings that underlie knee-jerk responses to eliminativism—Can't you *see* that statue? Do you think we are *dreaming* that statues exist?—and to counter two different versions of the charge that eliminativism is contradictory.

I. Eliminativism: The Basic Idea

Let us start by contrasting the eliminativist's claim that there are no statues with the claim that there is no Bigfoot. (I'll focus on eliminativism about statues in this chapter; the points made apply to eliminativism about non-living macrophysical objects generally.) The claim that there is no Bigfoot, as it is typically understood, dictates a certain array of responses to a terrified camper who honestly claims to have seen a Sasquatch. We might suspect that she has hallucinated, been the victim of a prank, or perhaps mistaken something for Bigfoot, such as a bear, that, upon further inspection, the camper herself would agree is not Bigfoot.

But when someone claims to have seen a statue, I am *not* likely to suspect that he has hallucinated, been the victim of a prank, or mistaken something for a statue that, upon further inspection, he would agree was not really a statue after all. Such occurrences are genuine possibilities. For example, my mother once mistook a very still alligator for a statue (oops). But they are exceptions, not the rule. The rule is that one's 'seemingly seeing a statue' is caused—in a non-hallucinatory, non-prankish way—by things *arranged statuewise*.[1] And further inspection of those things would not lead one to deny one has really seen a statue (cf. §II).

[1] The locution 'arranged statuewise' is inspired by van Inwagen (1990: 109), though my explication of such expressions differs from his.

In much of what follows, I'll make claims about *atoms* arranged statuewise. I have in mind here the atoms of physics, not Democritus. For there is no need to build a commitment to (or, for that matter, against) simples into eliminativism (cf. Chapter 4, §VI). Then again, there is no need to build in a commitment to the atoms of physics either. So consider my claims about the atoms of physics to be useful but expendable. Such claims are really placeholders for claims about whatever microscopic entities are actually down there.

Obviously, that there are atoms arranged statuewise does not mean that there are atoms arranged so as to compose a statue. For if it did, eliminativism about statues would entail that there are no atoms arranged statuewise. To get a rough-and-ready understanding of what it does mean, imagine the eliminativist and the folk ontologist standing in front of an alleged statue, allegedly composed of atoms. Imagine further that the eliminativist wants to explain her position to the folk ontologist. She can do so by saying that those atoms—the ones in front of her—do not really compose a statue. But to do that, she must be able to refer to them. And she can refer to them with 'the atoms arranged statuewise'.

Let us go beyond this rough-and-ready understanding of 'arranged statuewise'. The first step is to consider the following thesis: atoms often compose statues; but whether atoms compose a statue does not supervene on *anything* about those atoms except, trivially, that they compose a statue. This thesis has the following result. There is a possible situation in which atoms compose a statue. There is a second possible situation in which atoms have all the same properties and stand in all the same relations (both to each other—save the superadded relation of *composing a statue*—and also to everything else in the world) as in the first situation but fail to compose a statue.

The above thesis is unattractive. So I think noneliminativist philosophers should assume that, if atoms compose a statue, then they do so because of something *else* about

them, something on which statue-composition non-trivially supervenes. To be a bit more precise, the sort of supervenience I have in mind here is both *global* and *microphysical*.[2] My claim, then, is that philosophers who believe in statues should assume that worlds that are exactly alike at the microphysical level are exactly alike with respect to the existence (and qualities) of statues.

Of course, one could also add that worlds alike in other ways are alike with respect to statues. For example, one could add that worlds alike with respect to statue-shaped lumps are alike with respect to statues. Such additions are not of concern in this section. All that matters here is that, if statues existed, then whether certain atoms compose a statue would nontrivially supervene on their features and the relations they stand in to each other and to all the other microscopica in the world.

Now consider the following:

> Atoms are *arranged statuewise* if and only if they both have the properties and also stand in the relations to microscopica upon which, if statues existed, those atoms' *composing a statue* would non-trivially supervene.[3]

From the perspective of the folk ontologist, there should be no complaints about this account of 'arranged statuewise'. It is unexceptionable.

[2] Or mostly microphysical. I'll argue in Chs. 4 and 6 that the existence of humans with conscious mental properties does not supervene on the microphysical. So if the global supervenience base for statue-composition includes conscious human (e.g. artistic) intentions, then that base includes the microphysical *and* whatever is required for the mental. I'll ignore this qualification in what follows.

[3] This account of 'atoms arranged statuewise' indicates how to understand similar expressions used throughout the book. For example, to define 'atoms arranged *baseballwise* ', start with this account and make these substitutions: 'baseballwise' for 'statuewise', 'if baseballs existed' for 'if statues existed', and 'composing a baseball' for 'composing a statue'. To define '*things* arranged statuewise', substitute 'things' for 'atoms' in this account. And so on.

From the perspective of the eliminativist, however, the account is not quite perfect. Its imperfection is *not* that it implies that statue-composition follows from statuewise arrangement. The account does not imply this. (It implies only that if statues existed, then atoms arranged statuewise would compose a statue.) Rather, its blemish is that, from the perspective of the eliminativist, this account includes a counterpossible conditional whose antecedent is 'if statues existed'.

Only the most dogmatic metaphysician refuses to entertain seriously what she takes to be counterpossibles (cf. Chapter 2, §V). For so to refuse is to refuse even to consider the strengths (or weaknesses) of theories one believes to be necessarily false. Enemies of temporal parts, for example, should recognize that—*per impossibile*!—were four-dimensionalism true, there would be otherwise unavailable solutions to the Ship of Theseus. And enemies of temporal parts, in recognizing this, should do more than attribute vacuous truth to a counterpossible. In a similar spirit, the eliminativist should affirm the conditional embedded in the above account. And the eliminativist can then accept that account of 'arranged statuewise'.[4] (See Sider's 1999*a*: 339–40 defence of using counterpossibles in a similar way.)

Nevertheless, I concede that my opponents have in hand an account of 'arranged statuewise' that is, by their lights, better than the one I have. Such are the burdens of hospitality: giving the best to one's guests, reserving the least for oneself. Moreover, in the current dialectical context, what is most important is that eliminativism's opponents have a crystal-clear account of 'arranged statuewise'. For presumably eliminativism's defenders will not be tempted to quibble pedantically over this issue, freely admitting (what is surely

[4] The folk ontologist who *denies* that whether atoms compose a statue non-trivially supervenes should accept the account in a similar way, interpreting it to include a counterpossible with the antecedent 'if statues exist and statue-composition is non-trivially supervenient'.

true) that they get the idea behind 'arranged statuewise' well enough.

Some might want to exchange my account of 'arranged statuewise' for a 'fictionalist' account. The fictionalist account says that atoms are arranged statuewise just in case, according to the 'folk-ontological fiction', they have properties and stand in microscopic relations upon which their composing a statue supervenes. This no more requires counterpossibles, so the suggestion goes, than does the claim that, according to Leibniz's monadology, the table is really a colony of souls (see Rosen 1990: 331).

I can see why some who are uneasy with counterpossibles might find the fictionalist account attractive. But note that uneasiness with counterpossibles, and so the motivation for 'going fictionalist', will afflict only those who reject folk ontology. From the perspective of the folk ontologist, my account of 'arranged statuewise' does *not* invoke a counterpossible. And, as noted above, it is the folk ontologist I most want that account to please.

And I suspect that the folk-ontological fiction—or, to make this alternative account of 'arranged statuewise' an option for folk ontologists, the folk-ontological *story*—says little or nothing about the features of the microscopic entities upon which, if statues existed, their composing a statue would non-trivially supervene. In support of this, note that the relevant microscopica might turn out not to be atoms or indeed anything for which we currently have a name. On the other hand, the question of what is 'true according to a fiction (or story)' is a difficult one (see e.g. Lewis 1978). And so perhaps an account of 'arranged statuewise' in terms of the folk fiction (or story) can do the job. If so, I commend it as an option worth taking seriously.

Indeed, in filling out my own account of 'arranged statuewise', I am inclined to borrow a page from fictionalism. For I think that—given eliminativism—the folk concept of *statue*

(and so the folk fiction about statues) plays a crucial role in grounding the non-vacuous truth of the relevant counterpossible and so in fixing which atoms are arranged statuewise. That is, what the folk mean by 'statue' is part of what makes it the case that certain atomic features are those upon which, if there were statues, statue composition would supervene.[5] More about this later.

One might accept my account of being arranged statuewise, yet claim that that account could be understood only by one who already knew what statues were supposed to be. This claim, even if true, is no objection to my ontology. For comparison, imagine the philosopher who says that 'smiles' are not entities on people's faces, who says that there is no object such that it is a smile. But this philosopher is quick to add that people smile. Smiling is something they do. Jay Rosenberg puts this sort of idea in the following way:

> For a person to wear a warm, welcoming smile is not for him to stand in a quasi-sartorial relation to an independent entity, distinct from and somehow supervenient on his curved lips, but simply for him to smile, warmly and welcomingly. (1993: 701)

These claims seem quite sensible. But all that matters for our purposes is that the following, even if true, would be a bad reason to reject these sensible claims: to understand what it is to smile, one must first know what smiles are supposed to be. Likewise, no one should reject eliminativism on the grounds that, to understand what it is to be arranged statuewise, one must first know what statues are supposed to be.

In fact, the concept of *statue*'s being epistemically prior to the concept of *being arranged statuewise* fits nicely with my

[5] My account of 'arranged statuewise' does not, all by itself, entail that the folk concept of *statue* plays a role in fixing whether atoms are arranged statuewise. Folk ontologists could deny that the folk concept plays such a role, insisting instead that the conditional in my account is made true by the existence of statues and the appropriate facts about supervenience, all of which are independent of folk concepts.

overall account. After all, I say that the folk concept of *statue* plays a role in determining which atomic arrangements are statuewise. I would even go so far as to say that if *being arranged statuewise* were not derivative upon folk-ontological concepts in the ways noted here, something would be amiss. For the whole point of introducing statuewise arrangements is to show that eliminativism shares important common ground—the existence of atoms arranged statuewise—with folk ontology.

This common ground implies that those who think they see statues are, although mistaken, neither hallucinating nor the victims of a prankster. This common ground, as we shall see, implies that eliminativism is not a straightforwardly empirical thesis of the sort that could be settled simply by pointing out a statue. And I shall argue that these and other fruits of this common ground make eliminativism less radical than it might initially seem.

II. Eliminativism: Not as Bad as You Might Think

Let us depart from both folk ontology and eliminativism for just a moment. Let us consider, instead, the claim that the atoms arranged my-neighbour's-dogwise and the-top-half-of-the-tree-in-my-backyardwise compose an object. There are legitimate ways to try to defend this claim. The most obvious involve philosophical arguments for unrestricted (universal) composition, the thesis that any two things compose something.[6] But it won't do to defend this claim with nothing more than 'I can *just see* the object composed of the atoms arranged dog-and-treetopwise'. Part of why this won't do, presumably,

[6] Defenders of unrestricted composition include Cartwright (1975), Leonard and Goodman (1940), Lewis (1986*a*, 212–13), and Sider (1997).

is that one's visual evidence would be the same *whether or not* those atoms composed something. Because this sort of defence won't do, I'll say debates over whether the dog-cum-treetop exists are not 'straightforwardly empirical'.

Whether atoms arranged my-neighbour's-dogwise and the-top-half-of-the-tree-in-my-backyardwise compose something is not a straightforwardly empirical question. By the same token, whether atoms arranged statuewise compose something (a statue) is not straightforwardly empirical. In part this is because, as with the dog&treetop, my visual evidence would be the same whether or not the atoms arranged statuewise composed something.

The analogy with the atoms arranged treetopwise and my-neighbour's-dogwise supports the claim that whether atoms arranged statuewise compose a statue is not straightforwardly empirical. Here is more support for that claim. The fundamental question is not so much whether some particular alleged statue exists. That question might—sceptical scenarios aside—seem to be a matter of just looking and seeing. The issue is rather whether, in general, atoms arranged statuewise compose a statue. And whether or not they do so will be a matter of necessity. But it seems that this question of metaphysical necessity cannot be decided, one way or the other, simply by a trip to the museum or a ride down Monument Avenue. It must be decided on philosophical grounds.

Atoms arranged statuewise cause the visual and other sensations the folk ontologist thinks are caused by statues. Indeed, atoms arranged statuewise can do just about anything normally attributed to statues. They can, for example, be purchased at auction and serve as landmarks. As a result, eliminativism need force no revision of our everyday practices like buying 'statues' or relying on 'statues' in giving directions. (In fact, I shall argue in Chapter Seven (§III) that eliminativism better accommodates some of our practices than does folk ontology.)

Moreover, eliminativism can even allow everyday statue-talk to remain largely unchanged. To begin to see this, consider whether 'the Crew of the USS *Enterprise*' is a plural referring expression—akin to 'Locke, Berkeley, and Hume'—or, instead, the name of a single large object with each crew member as a proper part. Note, in fact, that there are two questions here. First, there is the semantic question of what 'the Crew of the USS *Enterprise*' is supposed to mean. Second, there is the metaphysical question of whether there really is a big physical object that has all and only the crew members as its parts (at one level of decomposition[7]), a scattered object that weighs as much as the sum of the weights of those people taken individually.

I am not sure how to answer the first question. But, I say, the answer to the second question is 'no'. Some philosophers would disagree. No matter. The point here—in this section of this chapter—is not to settle either the metaphysical or the semantic dispute surrounding 'the Crew of the USS *Enterprise*'. It is, rather, that such disputes are neither here nor there with respect to everyday uses of 'the Crew of the USS *Enterprise*'. 'The Crew of the USS *Enterprise*' will continue to perform its ordinary duties regardless of how or whether the semantic and metaphysical disputes get settled.

Similarly, everyday sorts of claims like 'there is a statue aboard ship' can be useful and appropriate whether or not there are, in addition to atoms arranged statuewise, statues. Thus for practical purposes, the bulk of everyday uses of 'statue'—as well as, for example, 'Michelangelo's *David*' and '*the Statue of Liberty*'—can remain unchanged, even if the claims such uses express are false because eliminativism is true.

[7] Intuitively, an object's 'parts at one level of decomposition' are parts of that object that do not overlap and that, collectively, fill the whole region the object fills. For example, an object's parts at one level of decomposition might be its atoms and, at another, its elementary particles.

The fact that the eliminativist insists that there are atoms arranged statuewise suggests another way—besides highlighting eliminativism's non-straightforwardly-empirical and non-practical nature—that eliminativism is not as radical as it might first seem. Some (presumably non-eliminativist) philosophers say that there is nothing more to the material world than microscopic entities arranged in certain ways. They even think of this position as part of our ordinary, or at least of our 'scientific', outlook on the world. Thus Richard Swinburne simply asserts, as if it were obvious, that:

> there is nothing more to large-scale material objects except the fundamental particles and the relations they have to each other. (1995, 395)

And John Searle tells us:

> the basic intuition that underlies the concept of reductionism seems to be the idea that certain things might be shown to be nothing but certain other sorts of things. The most important form of reduction is ontological reduction. It is the form in which objects of certain types can be shown to consist in nothing but objects of other types . . . This form is clearly important in the history of science. For example, material objects in general can be shown to be nothing but collections of molecules . . . (1992: 112–113)

And consider reductionism as characterized (but not endorsed) by Hilary Kornblith:

> the inventory of microphysics is in some important sense complete: once we are done specifying all of the microphysical things there are, we have specified all of the things there are. (1993: 53)

It is hard to know how exactly to interpret 'nothing more to' or 'nothing but'. But eliminativism provides one ready interpretation. Saying there is *nothing more to* a statue than—or saying that a statue is *nothing but*—atoms interrelated in certain ways can readily be interpreted as meaning that, when it comes to alleged statues, there are really only atoms in statuesque

arrangement and nothing else at all. On this interpretation, an interpretation buttressed by Kornblith's comments, 'scientific reductionism' implies eliminativism. (But scientific reductionists, thus interpreted, go further than I do, eliminating all 'large-scale' material objects, including human organisms.)

The familiar sentiments noted by Kornblith, and endorsed by Swinburne and Searle and of course many others, do not amount unequivocally to eliminativism. And I am sure Swinburne and Searle would disavow eliminativist interpretations of their positions. (In § IV I shall quote Searle's distinctly non-eliminativist gloss on reductionism.) Nevertheless, it is worth noting that familiar reductionist sentiments resonate with eliminativism, making eliminativism somewhat more plausible and somewhat less foreign than it might otherwise be.

III. The Linguistic Charge of Contradiction

If eliminativism is to get a hearing, it is important to establish that it is not empirically falsified simply by presenting, à la G. E. Moore, one (alleged!) statue and then another. In establishing this, I relied on the fact that, although there are no statues, there are atoms arranged statuewise. I also emphasized that such atoms do most or even all of the everyday work allegedly done by statues. Moreover, I said that eliminativism coheres with familiar reductionist claims, claims like a statue is 'nothing but' atoms arranged statuewise, there is 'nothing more to' a statue than its constituent atoms, and the inventory of microphysics is 'complete'.

All of these considerations are intended to support eliminativism. But they might make atoms arranged statuewise sound suspiciously like statues. And this is bad news for the eliminativist. For if atoms arranged statuewise are statues, then there

is—to say the least—a tension between the claim that there are no statues and the claim that there are atoms arranged statuewise. There are, in fact, two distinct ways to argue that these claims actually *contradict* each other. I shall develop, and respond to, one such argument in this section and another in the next.

'There are married bachelors' is not explicitly formally contradictory, but it is contradictory in some quite straightforward sense. And one might object that 'there are atoms arranged statuewise but no statues' is contradictory in the same way. For as 'bachelor' means someone who is, among other things, unmarried, so—the objector insists—'there are [composite] statues' *just means* that there are some things arranged statuewise. Because of its contradictory nature, we should not take seriously an ontology according to which there are married bachelors. Likewise, this objection concludes, we should not take seriously the eliminativist's ontology with its atoms arranged statuewise but no statues (see Hirsch 1993). This—the 'linguistic' charge of contradiction—is the first of the two arguments for the claim that eliminativism is contradictory.

The linguistic charge has a faulty foundation. 'There are statues' does *not* mean only that there are some things arranged statuewise, in the sense of 'arranged statuewise' at issue in this book. Again, 'there are statues' does not mean that there are things that both have the properties and also stand in the relations to microscopica upon which, if there were statues, their composing a statue would non-trivially supervene. This is simply not a plausible claim about ordinary meaning.

Some might cry 'Paradox of Analysis!' They might say that the above claim about ordinary meaning—although it does not reflect what most people would initially *say* is meant by 'there are statues'—nonetheless provides an analysis of what is meant by 'there are statues'. And, they might conclude, this makes eliminativism contradictory. In reply, I recognize that genuine

analyses can be surprising. But they cannot be circular. Thus 'there are statues' cannot be analyzed in terms of how things would be if, among other things, there were statues. But this is just the sort of 'analysis' our objectors are here proposing.

One could reply that the proposed analysis equivocates on 'there are statues': in the analysans it means there are *really* statues, whereas in the analysandum it means merely that there are only things arranged statuewise. This cures the analysis of circularity. But, obviously, the linguistic charger cannot take this medicine; it requires the very distinction she finds contradictory, a distinction between there really being statues and there being only things arranged statuewise. (On the other hand, the *eliminativist* who (rejects the linguistic charge and) endorses the above analysis of the ordinary meaning of 'there are statues' could take it to be non-circular in the way just suggested.)

At any rate, I say the linguistic charge rests on false claims about the ordinary meaning of 'there are statues'. Now perhaps there is a more subtle objection in the area, an objection not predicated on false or unacceptably circular claims about what is ordinarily meant by 'there are statues'. But note that the 'more subtle' objection cannot merely be, for example, that there being atoms arranged statuewise is *sufficient* for the truth of—or provides one set of truth conditions for—what is ordinarily meant by 'there are statues'. This is merely to deny eliminativism, which is not at all the same thing as showing it to be contradictory.

It is difficult to see how the linguistic charge can avoid false claims about meaning and synonymy while also preserving its lofty status as a charge of contradiction. For it seems that once it abandons the false claims about meaning, it must be downgraded to the bald assertion that, because statues exist, atoms arranged statuewise always compose a statue. Nevertheless, there may be some accounts of meaning that allow the linguistic charge to avoid these pitfalls. So I want to respond to

the linguistic charge supposing, for the sake of argument, that 'there are statues' ordinarily means only that there are atoms arranged statuewise.

Given this supposition, the eliminativist should insist that sometimes—such as when it is denied by eliminativists—the sentence 'there are statues' means something else. Sometimes, according to the eliminativist, 'there are statues' means that atoms arranged statuewise stand in the relation of *composing something* one to another. Sometimes 'there are statues' means that there is, in addition to various atoms in statuesque arrangement, some much bigger object—with a mass, centre of gravity, and so on—that has each of those atoms as a part. And when this is the meaning of 'there are statues', says the eliminativist, 'there are statues' is false (cf. van Inwagen 1993).

To better understand this response, compare it with the following claims. 'There is a crew aboard ship' is true of any fully manned ship (pretend there are ships), if it means only that there are many people aboard, performing certain assigned tasks. But 'there is a crew aboard ship' is false if it means that those people stand in the relation of *composing something* one to another. It is false, that is, if it means that on the ship there is a big scattered object composed of those people, an object with a mass many times that of any person and a constantly shifting and hard-to-locate centre of gravity.

This explanation of eliminativism in terms of *composing something* should be helpful to some, especially when compared to parallel claims about crews. But it will do little to placate many originally inclined to endorse the linguistic charge. For many moved by the linguistic charge in the first place will think that 'there are atoms that compose something' means only that there are atoms arranged in one or another of a variety of ways, including statuewise. And so they are likely to charge that one contradiction has been exchanged for another.

So let us try again. Suppose there are *n* atoms arranged statuewise in a room. 'There is a statue in the room' in its

allegedly ordinary sense implies only that there are *n* things in the room—the atoms themselves—as big as, or bigger than, an atom. 'There is a statue in the room', in the sense in which it is denied by the eliminativist, implies that there are at least *n* + 1 things in that room as big as, or bigger than, an atom.

For comparison, note that one reading of 'there is a forty-member crew aboard ship' implies only that there are forty physical objects on the ship that have a heart among their proper parts. A second reading implies that there are at least forty-one physical objects on the ship that have a heart (i.e. have at least one heart) among their proper parts: the members and the crew. The first reading parallels the allegedly ordinary meaning of 'there are statues'. The second parallels the 'other' meaning of that sentence, the one the eliminativist says is false.

These comments should be enough to explain how the eliminativist's denial of 'there are statues' does not contradict the claim that there are atoms arranged statuewise. One could, however, object that these comments are themselves contradictory. One might claim that 'there are *n* + 1 things in the room as big as, or bigger than, an atom' means that there are *n* at-least-atom-sized things in the room arranged in one or another of various ways, including statuewise. On the one hand, I think this shows we have successfully responded to the linguistic charge. For this is not a plausible claim about ordinary meaning. On the other hand, it shows how difficult it is to respond to that charge. For it shows that there will always be a (dubious) claim about meaning the linguistic charger can make that, if true, would render any direct explanation of eliminativism contradictory.

Let us try an indirect explanation, by way of an analogy. Suppose that I exist and person P exists. And suppose, for the sake of argument, that there is nothing composed of all and only the atoms arranged my-left-earwise and P's-nosewise. In asking us to make this second supposition, I am asking us to suppose that unrestricted composition is false. Making this

supposition does not require that unrestricted composition actually *be* false, or even possibly false. All it requires, given the aims of the analogy I'm developing, is that the denial of unrestricted composition is not contradictory *in the sense at issue in the linguistic charge*. This requirement can be met. For we should all agree that our ordinary ways of speaking do not render the denial of unrestricted composition contradictory. (Note: Defenders of unrestricted composition *defend* it. They do not think that 'any two objects compose an object' is akin to 'there are no married bachelors', which needs no defence.[8])

Again, suppose that there is nothing composed of all and only the atoms arranged my-left-earwise and P's-nosewise. Now imagine a world just like ours with respect to which things exist. So P and I exist there. And given what we are supposing about our world, there is nothing in that world composed of the atoms of my left ear and P's nose. In the world we are imagining, moreover, everyone speaks English. But in the dictionaries of that world there is an entry for 'slithy tove' which reads: 'an object composed of the atoms of Merricks's left ear and the atoms of P's nose'. Let us add that this definition of 'slithy tove' is widely known in that world and the expression itself widely used.

I assume that a philosopher in this imagined world could grant, for the sake of argument, that 'there is a slithy tove' ordinarily means only that there are atoms arranged my-earwise and P's-nosewise. He could therefore grant that 'there is a slithy tove' is ordinarily true. Yet it also seems—this is a crucial assumption—that this philosopher should be able to state *the truth*. He should be able say that *there is nothing in existence* that is the referent of 'slithy tove'. He should be able to say truly that the atoms of my left ear and of P's nose do not, in concert, compose any object at all. (None of this is to deny that our

[8] Those who think a whole is identical with its parts might say that unrestricted composition is 'ontologically innocent' (Lewis 1991: 81) and so not in need of defence. Below (§IV) I argue that a whole is not identical with its parts.

philosopher might have a difficult time explaining the truth. It might take him a whole chapter to get the idea across.)

I say that, as it is with our imagined philosopher and 'there is a slithy tove', so it is with eliminativists and 'there is a statue'. One might deny the situations are analogous. That is to deny eliminativism. But all that matters here is that to say that the situations are analogous—and so to affirm eliminativism—is not like saying 'there are married bachelors'. Thus eliminativism has been explained in a way not open to the linguistic charge of contradiction.

There is a second point. The linguistic charge could be levelled, *mutatis mutandis,* by our philosopher's worldmates against him when he says 'there is no slithy tove, although there are atoms arranged Merricks's-left-earwise-and-P's-nosewise'. But in the imagined world the linguistic charge is mistaken. For rightly understood, what the imagined philosopher is saying is true. I think this shows that the linguistic charge in our own world, applied to eliminativism about statues, is mistaken as well.

Let me focus on what I think is the fundamental point at issue. I assume that there is an objective fact of the matter about what exists. And I think we use the apparatus of existential quantification—expressions like 'there is', 'there are', and 'exists'—to say what (we believe) objectively exists. But there is nothing magical about 'there is', 'there are', or 'exists'. We control them; they do not control us. So we can use these bits of language however we choose. Thus we could use them 'deviantly', to do something other than describe what (we believe) exists. For example, we could use 'there is an F' to mean we wish there were an F.

With this in mind, let's return to the linguistic charge. If the linguistic charge's assumption about the ordinary meaning of 'there is a statue' is correct, 'there is a statue' does *not* ordinarily mean that there is some *x,* such that *x* is a statue. It means, instead, that there are some things, none of which is a statue,

in certain arrangements. Thus if the linguistic charge's claim about ordinary meaning is correct, then 'there is' is used deviantly in ordinary occurrences of 'there is a statue'. Eliminativism has nothing to say about such deviancy. Eliminativism claims only that 'there is a statue' is false when 'there is' is being used as a legitimate and straightforward existential quantifier.

I think that reflecting upon deviant versus straightforward uses of 'there are', 'there is', and 'exists' supports my original response to the linguistic charge. That response was that 'there are statues' does *not* ordinarily mean that there are atoms arranged statuewise. For I think—although I don't really care—that most ordinary speakers do not use 'there are' deviantly when they say 'there are statues'. Contrary to the allegations of the linguistic charge, it seems clear that the most literal and straightforward meaning of 'there are statues' is also the ordinary meaning. So in what follows I shall assume that—if eliminativism is true—when the folk say 'there are statues', they say something false. (I shall return to this issue in Chapter 7, §I.)[9]

I have responded to the linguistic charge. I have explained why believing in atoms arranged statuewise while disbelieving in statues is not akin to believing in married bachelors. But I don't harbour any illusions. I know there are some who are still swayed, as a result of their philosophy of language, by the linguistic charge. They will find themselves unable even to understand the ontology I defend. Those same philosophers do not understand, or at least should not understand, ontologists who deny the existence of numbers, properties, or holes. This illustrates that the eliminativist is, with respect to

[9] Some might object that there is no single straightforward and literal meaning of 'exist'. I say more about this in Ch. Seven (§I). Here I note only that this objection does not, all by itself, imply that one of the allegedly many ordinary, literal, and straightforward meanings of 'statues exist' is that there are things, none of which is a statue, arranged in certain ways.

the linguistic charge, no worse off than any (other) nominalist. I don't know what more to say in defence of the discipline of ontology, other than to note that the two chapters that follow may help. For those chapters will highlight some of the substantive issues at stake in the debate over whether, in addition to atoms arranged statuewise, there really are any statues.

IV. The Metaphysical Charge of Contradiction

Recall that, in a passage quoted above (§II), Searle says that a material object is *nothing but* certain other sorts of things. Searle then—and here we go beyond previously quoted material—calls the 'nothing but' relation 'a peculiar form of the *identity* relation' (1992: 113; emphasis added). Similarly, David Armstrong says 'mereological wholes are *identical* with all their parts taken together' (1997: 12). Donald Baxter states 'the whole is many parts counted as one thing . . . there is no one thing distinct from each of the parts which is the whole' (1988: 578). David Lewis agrees that the parts of a thing are *identical* with the whole they compose.[10] Lewis glosses this view—he calls it 'composition as identity'—as the claim that a physical object is 'nothing over and above its parts' (1991: 80). This gloss brings to mind the familiar reductionist sentiments noted earlier in the chapter.

Suppose, if you can, that composition as identity is true.[11] Suppose also that statues exist composed of atoms. Then each

[10] But Lewis disagrees with the details of Baxter's position and makes distinctions Baxter does not. For Lewis says that there is a difference between the sort of identity that holds between a whole and its parts—he calls this the 'broadened' sense of identity—and the more familiar version of identity—what he calls the 'ordinary one-one' sense of identity. Lewis thinks these two senses of identity are analogous in significant ways (1991: 84–7).

[11] *Composition* as identity is not the view that *constitution* is identity. The former claims that a single object (e.g. a statue) is identical with the many parts (e.g. atoms) it comprises. The latter claims that a single object (e.g. a statue) is identical with a single object (e.g. a statue-shaped lump of clay) that 'constitutes' it.

statue is *identical with* some atoms (or others) arranged statuewise. And so one could charge that my ontology, with its atoms arranged statuewise but no statues, is in some way contradictory. In what way? In just the same way that an ontology that includes a certain object, but not something identical with that object, is contradictory. This—the 'metaphysical' charge of contradiction—is the second argument for the claim that eliminativism is contradictory.

Do not worry about this charge. Just reject composition as identity. One good reason to reject composition as identity is that it implies, obviously enough, that one thing (e.g. a whole) can be identical with many things (e.g. the whole's parts). But I think one of the most obvious facts about *identity* is that while it holds both one-one (John is identical with Mr Smith) and perhaps even many-many (John and Mary are identical with Mr Smith and Ms Jones), it never holds one-many.[12] (I follow Lewis 1991 in using 'one-one' to describe an ordinary binary relation, 'one-many' to describe a multigrade relation holding between one thing and many things, and 'many-many' to describe a multigrade relation holding between many things and many things.)

Identity cannot hold one-many. So composition as identity is false. Moreover, consider the following argument. Suppose that composition as identity is true. Suppose, then, that I really am identical with my constituent atoms $A_1 \ldots A_n$. And $A_1 \ldots A_n$ are identical with me. Identity is not temporary.[13] And so it follows that I am always identical with $A_1 \ldots A_n$ and they are always identical with me. Thus if composition as identity is true, the atoms that compose me have always composed me

[12] Let us grant, for the sake of argument and as a concession to my opponent, that identity can hold many-many. For if identity cannot hold many-many, then it cannot hold one-many—to see this, think about the transitivity of identity.

[13] Perhaps one could make sense of 'temporary identity' by way of temporal parts. (Some say that objects that share a temporal part are thereby temporarily identical.) Below I concede that temporal parts undermine the above argument.

and will always compose me. (I assume that if composition as identity is true, any atoms that are identical with me are thereby atoms that compose me.)

Composition as identity implies that I have not undergone, and will not undergo, change of atomic parts. More generally, it implies that no persisting object ever changes parts. This implication is false. So composition as identity is false.

Some will object that objects do not 'change parts' in the sense presupposed by my argument. But they will hasten to add—lest their position be incredible—that objects do change parts in another sense. To understand their objection, we must understand the thesis that objects perdure, that they are four-dimensional, that they have temporal parts. If I perdure, I am not 'wholly present' at any one moment in time. Rather, I am spread out over time much as a spatially extended object is spread out over space. And the perdurantist who believes in composition as identity will say that I am identical with all my parts, including my *temporal* parts. And, she will add, the four-dimensional Merricks never changes with respect to the parts he has. Thus, she will conclude, the above argument against composition as identity is unsound. (For more on perdurance and its rival, endurance, see Merricks 1994, 1995*a,* 1999*a, c,* and Chapter 2, §III.)

As just noted, the perdurantist denies that I change parts in the way required by my argument against composition as identity. Yet she will say that a temporal part of mine existing at one time has parts that a temporal part of mine existing at another lacks. Thus, she concludes, there is a straightforward sense in which I have different parts at different times. And this, she maintains, implies that there is a straightforward sense in which I experience change of parts, keeping her position on change from being incredible.

(The perdurantist blocks my argument against composition as identity by defending—and rendering credible—the claim that, in some important sense, objects do not change parts.

The charge that perdurance does not allow for real change is both shopworn and controversial. But the above shows that the perdurantist—at least if she believes in composition as identity—should positively insist that there is a sense in which objects do not change parts. Thus she should concede that there is some truth to the old charge. I think this tells against perdurance.)

At any rate, if objects do not perdure (but instead endure), then some objects do change parts in the sense presupposed by my argument against composition as identity. So if objects endure, then we must—in light of that argument—conclude that composition as identity is false.

There is a second argument against composition as identity, an argument that does not presuppose that objects endure. This second argument is a 'modal' version of the first. It starts by noting that Locke, Berkeley, and Hume are identical with Locke, Berkeley, and Hume. There is no possible world in which Locke, Berkeley, and Hume exist and even one of Locke, Berkeley, and Hume does not exist. And Locke, Berkeley, and Hume are identical with Locke, Berkeley, and Hume in every possible world.[14] Similarly, there is no possible world in which $O_1 \ldots O_n$ exist and even one of $O_1 \ldots O_n$ does not exist. And $O_1 \ldots O_n$ are identical with $O_1 \ldots O_n$ in every possible world.

Now suppose that O, the object composed of $O_1 \ldots O_n$, is identical with $O_1 \ldots O_n$. This, the fact that $O_1 \ldots O_n$ are identical with $O_1 \ldots O_n$ in every possible world, and the indiscernibility of identicals imply that O is identical with $O_1 \ldots O_n$ in every possible world.[15] Therefore, if composition as identity is

[14] That is, Locke, Berkeley, and Hume are identical with Locke, Berkeley, and Hume in every possible world *in which they exist.* In what follows, claims about an object's identity 'in all possible worlds' are to be understood as implicitly restricted to worlds in which the object exists.

[15] Lewis hedges on whether 'broadened' identity implies indiscernibility: 'even though the many and the one are the same portion of Reality, and the character of that portion is given once and for all whether we take it as many or take it as one, still we do not really have a generalized principle of indiscernibility of identicals. It does matter how you slice it—not to the character of what's described, of course, but to the form of the description' (1991: 87). My arguments require only the indiscernibility of the 'character' of the one and the many.

true, there is no world in which O exists but is not composed of $O_1 \ldots O_n$. So composition as identity implies that O—and, of course, every other composite object—must, in every world in which it exists, be composed of the parts that actually compose it. In other words, composition as identity entails mereological essentialism. But—here I make a fairly uncontroversial assumption—mereological essentialism is false. Therefore composition as identity is false. QED.

I feel entitled to the commonsense denial of mereological essentialism. And in presenting my 'temporal' argument, I helped myself to the bit of common sense that says objects change parts over time. But I reject the bit that says there are statues. Problem? No. I start off with all the common sense I can get, unloading it only when prompted to do so by the overall burden of argument. (More on this topic in the introduction to Chapter 4.) By the end of the book it should be clear why I abandon the bit of common sense that says there are statues. In my opinion, there are not equally compelling reasons to jettison the bit that rejects mereological essentialism (or the bit that says objects can change parts over time). I may disagree with most readers about which composite objects exist. But we can still agree that there is some composite object or other—such as, I say, a human organism—that does not have all its parts essentially.

Given the necessity of identity, the modal argument against composition as identity goes through. And I accept the necessity of identity. So I conclude that we have a sound argument for the falsity of composition as identity. Yet there are fans of contingent identity. For the sake of argument, I want to explore their options. For, surprisingly, contingent identity alone is not enough to undermine the argument. I'll explain

why by making use of a version of counterpart theory according to which the counterpart relation is *not* sortal-relative. No one defends this kind of counterpart theory. (At least not any more; it was defended in Lewis 1968.) But considering it is the best way to show that contingent identity alone cannot undermine the above argument.

The counterpart theorist says that Locke, Berkeley, and Hume can share a single counterpart in another world—a world in which there is, say, just one British Empiricist. Thus the counterpart theorist can insist that Locke, Berkeley, and Hume all exist in a world in which, so to speak, there is only one of them. In that world, by the counterpart theorist's lights, Locke is identical with Berkeley is identical with Hume. Yet that identity is contingent. In the actual world they are distinct.

Similarly, the counterpart theorist says that $O_1 \ldots O_n$ can exist in a world in which there are fewer than n of them. For in some world several of $O_1 \ldots O_n$ have the same counterpart and are therefore contingently identical. In this way, the defender of counterpart theory and composition as identity can maintain that O, the object composed of (and identical with) $O_1 \ldots O_n$, can exist in a world in which it has fewer parts than it actually has.

Now according to the counterpart theorist-cum-devotee of composition as identity, the world in which O has fewer than n parts, fewer than it actually has, is *ex hypothesi* one in which it nevertheless has $O_1 \ldots O_n$ as parts. (In that world, however, some of $O_1 \ldots O_n$ are contingently identical with each other.) It is by saying this—by insisting that O's counterpart (or counterparts) in every world has (or have) counterparts of $O_1 \ldots O_n$ as parts—that she maintains that her view is consistent with composition as identity. But it is by saying this that she also commits herself to mereological essentialism. This shows that *mere* contingent identity is not enough to undermine the simple argument above for composition as identity's entailing mereological essentialism.

Contingent identity alone does not undermine that simple argument. But a species of counterpart theory—the species some philosophers actually defend—does. That species insists that objects do not have counterparts *simpliciter,* but rather only *qua* their being certain kinds of things.[16] Thus Locke *qua* influential Enlightenment figure might have a counterpart C in W, a figure that exemplifies a Lockean influence on the (counterpart of) the Enlightenment in W. But Locke *qua* descendant of ancestors $A_1 \ldots A_n$ might have a counterpart in W distinct from C, one whose ancestors are (counterparts of) $A_1 \ldots A_n$. So, although Locke *qua* Enlightenment figure is actually identical with Locke *qua* descendant, possibly—for example, in W—this identity does not hold.

Similarly, this species of counterpart theory says that O (that is, $O_1 \ldots O_n$) *qua* the many objects $O_1 \ldots O_n$ exists in other worlds only if counterparts of $O_1 \ldots O_n$ exist in those worlds. But O *qua*—for example—the single object named 'the Eiffel Tower', standing in Paris, and having shape S and mass M, exists in other worlds if those worlds contain the right sort of tower in (the counterpart of) Paris—even if they fail to contain counterparts of $O_1 \ldots O_n$. Thus, the counterpart theorist could grant that O is identical with $O_1 \ldots O_n$, yet note that O, *qua* something-other-than-$O_1 \ldots O_n$, is possibly not identical with $O_1 \ldots O_n$. As a result, O (*qua* something-other-than-$O_1 \ldots O_n$) could fail to have $O_1 \ldots O_n$ as parts.

So the species of counterpart theory according to which objects do not have counterparts *simpliciter* undermines my second argument against composition as identity. It does so by way of a *relativized* version of contingent identity (but not anything

[16] David Lewis (1971; 1986*a*: 248 ff.) and Allan Gibbard (1975) defend views along these lines. The rejection of counterparts *simpliciter* is more understandable given that, according to Lewis (but not Gibbard; see Gibbard 1975, n. 3), the counterpart relation is a similarity relation. It could be, for example, that A and B are more similar *qua* profession to each other than either is to C; B and C more similar *qua* gender to each other than either is to A; and questions of who is more similar *simpliciter* to whom ill formed.

like the 'relative identity' of Geach 1980). Indeed, the only way to undermine that argument is to endorse 'relativized' contingent identity. And I believe that sortal-relative counterpart theory provides the only way to make good sense of this sort of identity. So, I conclude, only those who endorse (sortal-relative) counterpart theory can sensibly resist the argument linking mereological essentialism to composition as identity.

But we should not endorse sortal-relative counterpart theory. This is because it implies that identity is, in the ways noted above, contingent and 'relativized'. Although sortalrelative counterpart theory makes the claim that identity has these features coherent, that claim still seems false. Moreover, many deny that counterpart theory can even satisfy the demand placed upon it by its very *raison d'être,* that of providing a compelling account of *de re* modality. Here is an example of the most familiar sort of objection along these lines: the existence of someone in another world who is a lot like me, but happier, is irrelevant to whether *I*—this very person—could have been happier; and it is irrelevant even if we call that other-worldly someone 'my counterpart' (see Kripke 1980: 45–9; Plantinga 1974: 108–20).

If, for whatever reason, we reject sortal-relative counterpart theory (as I do), we must accept that composition as identity implies mereological essentialism. And if we also reject mereological essentialism (as I do), we must conclude that composition as identity is false.[17] Similarly, if we believe (as I do) that objects endure and that at least some objects change parts over time, then we should reject composition as identity. Henceforth, I shall proceed on the assumption that composition as

[17] In the next chapter I say that even if there were objects (like masses) that have all their parts essentially, they would not be identical with their parts. It would be a mistake to object that, since my argument against composition as identity relied on the denial of mereological essentialism, composition could be identity in cases involving the mereologically invariant. For the conclusion of my argument is about composition *as such* : composition is not identity.

identity is false. I proceed thus because of the temporal and modal arguments just given and also because the claim that identity holds one-many seems contrary to the logic of identity.

Composition as identity is false. So every composite object is distinct from—i.e. not identical with—its parts. So every such object is something 'in addition to' its parts. For if we set out to take an inventory of all the objects in the universe, and included on our list only the parts of a composite object but not the object itself, we would have an *incomplete* inventory. It would be incomplete because there would be an object such that nothing identical to it is inventoried.

I don't know whether a composite object's being distinct from and in addition to its parts is consistent with 'reductionism' about physical objects. This is in large part because—once reductionism is distinguished both from composition as identity and from eliminativism—I don't know what it is supposed to be (see Merricks 1999*c,* §IV). But for our purposes, all that matters is that, no matter how 'reductionism' should be understood or what reductionism is consistent with, every object composed of atoms is not identical with, and so is something in addition to, those atoms.

Composition as identity is false. So even if there are statues, they are not identical with their constituent atoms (or, of course, with their proper parts at other levels of decomposition). Thus, even if atoms arranged statuewise do compose statues, ontologies that include atoms arranged statuewise but exclude statues are not contradictory. That is, they are not like ontologies that include an object but exclude something identical with it. The 'metaphysical' charge of contradiction cannot be made to stick.

V. Conclusion

Many disputed metaphysical theses are, if true at all, necessarily true. To establish that such a thesis is possibly true just is to establish that it is actually true. Thus establishing that a disputed metaphysical thesis is possibly true cannot be a prerequisite for presenting arguments for that thesis. Nevertheless, we should ask the metaphysician, before she presents her arguments, what thesis she purports to defend. After all, if we don't know what thesis she is defending, we cannot possibly judge whether her arguments adequately support it.

I have not tried to establish that eliminativism is possibly true. But I have tried to explain what eliminativism is. I have explained the eliminativist's thesis that there are atoms arranged statuewise but no statues. And I have countered two versions of the charge that the first half of that thesis contradicts the second. This should be enough to get us started. We should now understand eliminativism well enough to follow and evaluate arguments for its truth.

2 Considerations in Favour of Eliminativism

This chapter contains a number of arguments for eliminating various macrophysical entities alleged by folk ontologists to exist. These arguments, when taken together, constitute an attack on much of the folk inventory of material objects. And after we see how these arguments bear on persons in Chapter 5, we shall see that they present a compelling case for my overall ontology. But for now the point of this chapter—taken apart from Chapter 5—is rather modest. It is only to motivate eliminativism, to show why one might find it attractive.

I. The Water in the Pool

Consider the water in the swimming pool. Some—such as those who endorse unrestricted composition or those who believe in a kind of entity called 'a mass'—say that 'the water in the swimming pool' refers to a big material object. That object, they maintain, is shaped like a plaster cast of the swimming pool; it is about as tall as the pool is deep; it, like all material objects, has a mass and a centre of gravity. They will probably add that it has all its parts—or at least all its parts that

are properly called 'some water'—essentially (see e.g. Zimmerman 1995).

Why believe there is a big wet chunky thing that fits snugly into the pool? One might argue that it's the only thing keeping the high diver from crippling herself on the pool's cement bottom. Or one might say that we can just *see* 'the water'. Or one might suggest that 'the water' plays an integral role in fixing the truth value of sentences like 'the water in the pool today was not in the pool last year'.

We can easily resist these reasons. For suppose, instead of any one big chunky thing fitting snugly into the pool, there were only many, many H_2O molecules. (Composition as identity is false (Chapter 1, §IV). So the many H_2O molecules cannot be identical with the single object that is the water in the pool.) One could hold that those molecules prevent the diver from striking bottom. And one could insist that those molecules cause our visual sensation of seeing 'the water'. And suppose those molecules, which are in the pool today, were not there last year; that would make the following true: 'the water in the pool today was not in the pool last year'.

The point—in this section of this chapter—is not that there is no single big object, the water in the pool. The point is, rather, that 'eliminating' such an alleged object is not particularly radical. Perhaps folk ontology even favours this sort of eliminativism. Or perhaps folk ontology does not even raise, and so does not answer, the question of whether such an eliminativism is correct. But, whatever exactly we say here, I think it should be clear that the claim that there is no big chunky material object—the water—wedged into the pool is neither striking nor bizarre nor radical.

The fact that eliminating 'the water' is intuitively somewhat plausible, even before it is argued for, should lend at least some plausibility to eliminating statues. In so far as the folk ontologist is inclined to accept that there is no single big chunky object such as the water in the pool, but instead just H_2O

molecules suitably located and arranged, she is under some pressure to accept eliminativism about statues. After all, the water in the pool seems at least somewhat relevantly analogous to a statue.

We could further motivate this line of argument—that eliminativism is innocent by association—by considering other cases. We could return to an example from the first chapter, the Crew of the *Enterprise*. Or we could discuss the sand on the beach, the cheese in the kitchen, the rainforest in Brazil, or the furniture in the store. But I think that the point, a fairly weak and defeasible one, should already be pretty clear.

One might object that human organisms (and atoms) are analogous to the water in the pool. Thus one might argue that we have here no support for my overall ontology, no support for eliminating statues but embracing humans (and atoms). I shall respond to this objection, along with all other objections in this chapter that focus on the positive part of my ontology, in Chapter 5. Only then will we have the tools to respond adequately.

II. The Sorites Game

Let us suppose, for *reductio,* that the atoms arranged Davidwise compose Michelangelo's *David*. Now imagine that God agrees to play 'the Sorites Game' with that statue. We annihilate *David*'s atoms, one at a time—playing the Game is delicate work—and after each annihilation ask God whether *David* still exists (cf. Heller 1996). After the first, God would say, presumably, 'yes, David still exists'. We then annihilate a second atom and ask the same question. Again God would presumably say 'yes'. At some point in this process God will shift from saying 'yes' to saying 'no', thus showing that a single annihilation takes *David* from determinately existing to determinately not

existing. That is, God will do so unless one of the following two positions is true.

The first position is that a single annihilation during the Game would make it metaphysically vague whether *David* exists. But some claim that metaphysical vagueness is simply unintelligible (Russell 1923; Dummett 1975; and Lewis 1986*a*: 212). Moreover, metaphysical vagueness of the sort at issue here entails metaphysically vague identity. For if it were metaphysically vague whether *David* survives a certain move in the Game, then it would be metaphysically vague whether the object existing after that move is identical with *David*. But many take this sort of vague identity to lead straight to contradiction (see Evans, 1978).

There is a second way to avoid the conclusion that, at some point in the Game, a single annihilation would take *David* from determinately existing to determinately failing to exist. One could say that at the start of the Game there are many distinct objects, all of which are *equally good candidates* for being the referent of '*David*'; at some points in the Game, some, but not all, of those candidates exist; at those points, '*David* exists' is neither determinately true nor determinately false. This position does not require metaphysical vagueness. But it does require 'co-location' at a time of many *David*-candidates. In the next section I'll argue that we should reject such colocation. For now, I'll simply assume it should be rejected.

I deny that there are co-located multiple *David*-candidates. And I reject metaphysical vagueness. So I conclude that, if *David* exists at the start of the Game, then there is a point during the Game at which annihilating a single atom takes *David* from determinate existence to determinate non-existence. Thus we should conclude that *David* does not exist at the start of the Game. For the result that a single annihilation could make 'all the difference' in *David*'s existence is unacceptable.

This result should be unacceptable even to defenders of epistemic accounts of vagueness. The most developed

epistemic accounts, those of Williamson (1994) and Sorensen (1988), ground vagueness in our ignorance of the meanings or extensions of words. (Heller 2000 instructively calls this the 'episte-linguistic approach'. See also Merricks 2001) I believe that every plausible linguistic account of vagueness—and thus epistemic versions thereof—requires there to be, in some sense, equally good candidates either for what a vague predicate might mean or for what a vague name might refer to. For example, if the vagueness of baldness is rooted in 'bald', then there are many properties that are, in some sense, equally good candidates for being expressed by 'bald'.

But if *David* exists, then—given the assumption noted above and to be defended in the next section—there are not multiple equally good candidates for being referred to by '*David*'. So linguistic accounts of *David*'s vague persistence cannot find a foothold.[18] Episte-linguistic accounts, being linguistic accounts, inherit this problem. They have the further problem that, if *David* exists, then—given that there is not more than one *David*-candidate—it is *obvious* which object is referred to by '*David*'. And so we lack the requisite ignorance.

So if *David* exists, the Sorites Game has unacceptable results. Suppose instead that eliminativism is true. Suppose, then, that *David* does not exist at the start of the Game. We try to play the Game. We annihilate one of the atoms arranged Davidwise and ask God whether *David* still exists. God replies that once we know that a particular atom has been annihilated, and the others left in place, *we know everything there is to know.*

[18] I suppose one might claim that the candidates are not co-located, but instead overlap almost completely. This is not a standard view about candidates. And it has problems. For it implies that the candidates, alike but for the odd atom here or there, differ non-trivially in persistence conditions. Moreover, such a view of *David* will—unless it is objectionably *ad hoc* —fall out of a general metaphysics of material objects. But such a metaphysics implies that many humans that overlap nearly completely are now wearing my shirt, each of whom feels dead certain (and I would know) that he is the only one here. Unacceptable.

There is no further fact about whether *David* persisted throughout the episode, for *David* was never there to begin with. Most importantly, there cannot be a further fact to the effect that a particular annihilation took *David* from determinately existing to determinately not existing.

Eliminativism implies that there cannot be such a further fact. This is a more plausible result, I submit, than the result that there is such a further fact, forever hidden—absent Divine Revelation—from view. And it is more plausible than the Game's inducing metaphysical vagueness. And it is more plausible, for reasons I'll defend in the next section, than the co-location of many *David*-candidates. Because (so I say) eliminativism offers the most plausible understanding of what occurs during the Sorites Game, that game issues in a good reason to endorse eliminativism.

Some might think that just so long as *David* is 'reduced to' its constituent atoms, then knowing what happens to all the atoms is knowing all the facts (cf. Parfit 1984: 231–43). But unless 'reductionism' is a species of eliminativism, this is clearly mistaken. For suppose eliminativism is false. So *David* exists at the start of the Game. It then follows, whether or not 'reductionism' is true, that there will be some point during the Game at which we are in the dark as to whether it is true that *David* exists or instead something other than true (false, indeterminate, something). That is, there will be a point at which we are thus in the dark even though we know what has happened to all the atoms.

One might object that there is no uniquely closest world where we play the Sorites Game. And, one might continue, God's first 'no' answers occur at different points while playing the Game in different equally close worlds. Thus—contrary to the assumption underlying the above argument (so one might object)—there is no atom such that God would say 'no' for the first time, were that atom annihilated. (When comparing the Sorites Game across worlds, assume that in each world it is

played with the same atoms, occupying the same positions and annihilated in the same order.)

This objection takes aim at the wrong target. And so this objection is beside the point. For the point is not that there is some particular atom A such that it is the 'crucial' one. Rather, the point is that if we were to play the Game with *David,* then *David* would pass from determinately existing to determinately not existing upon the annihilation of a single atom. This is untouched by the above objection.

Moreover, we can—if we are so inclined—defend a point similar to the one the objection just noted targets, and can defend that point even while granting the objection itself. For, assuming *David* exists, God says 'yes' in all of the closest worlds when the first atom is annihilated. And in all of the closest worlds God says 'no' before all are annihilated. So it follows that there is an atom A whose annihilation, within some of the closest worlds, prompts God to say 'no' after saying 'yes' up to that point. That is, it follows that there is an atom whose annihilation *might* make all the difference between 'yes' and 'no' (see Lewis 1973*a, b*). It would be better to deny there is such an atom. And once we've set aside colocation and metaphysical vagueness, eliminativism is required to deny it.

Suppose we play a version of the Sorites Game asking God, after each annihilation, whether there are atoms *arranged Davidwise*. This version, one might object, reintroduces all the same problems attendant upon the original version of the Game. And thus, our objector concludes, no progress is made even if eliminativism is true, even if we exchange *David* for only atoms arranged Davidwise.

Atoms are arranged Davidwise only if they are such that, if *David* existed, they would compose *David*. The eliminativist says this conditional is a counterpossible. Part of what makes it non-vacuously true are the intrinsic and relational features of the atoms thus arranged. And part of what makes it non-vacuously

true, given eliminativism, is what the folk mean by '*David*' (Chapter 1 §I).

What the folk mean by '*David*' plays a role in whether atoms are arranged Davidwise. With this in mind, suppose we play the version of the Game in which we ask God whether there are still atoms arranged Davidwise. God's response need not, on the basis of a single annihilation, ever go from an unqualified 'yes' to an unqualified 'no'. God could say, at a certain point in the Game, that because what the folk mean by '*David*' is not maximally precise the following is true. There is no determinate fact of the matter about whether if '*David* exists' had been true at the start of the Game, then 'the atoms, given their features, compose *David*' would have been true at that point.

That is, God could say that at that point there is no determinate fact of the matter about whether 'those atoms are arranged Davidwise' is true. Thus we have the desired vagueness in the Game *without paying the price of either co-location or metaphysical vagueness*. Moreover, the vagueness we have is rooted in language, which seems desirable. And so I conclude that the results of playing the Sorites Game with atoms *arranged Davidwise* are far less problematic than the results of playing the Game with the statue *David*. (More on a related point in Chapter 7, §III.)

We can also easily accommodate an epistemic approach to the vagueness of whether atoms are arranged statuewise. For we can accommodate the following two claims. There is exactly one (perhaps disjunctive) arrangement such that it is picked out by 'arranged statuewise'. Because there are many very similar arrangements in the neighbourhood, we *cannot know*—save perhaps by playing the right version of the Game with God—which arrangement is the one so picked out. There is no need here for co-location of objects. We need only a plenitude of arrangements.

One might object that eliminativism does not ameliorate the real paradox underlying the Sorites Game. This is because,

the objection continues, eliminativism merely denies the existence of *some but not all* composite objects, while this paradox plagues composites of every sort. I respond to this objection in Chapter 5 (§II), where I address the objection that the Sorites Game raises the same problems for composite humans as it does for statues. I'll also say something there about microscopic composita.

III. The Statue and the Lump

Here is a familiar story:

> Consider a statue fashioned from a lump of clay. The statue is not identical with its constituent lump of clay because, among other things, the statue and the lump have different persistence conditions. The statue, but not the lump, could survive the loss of a few smallish bits of clay; the lump, but not the statue, could survive being squashed.

I find this story objectionable because I believe in neither statues nor lumps. But many other philosophers—many who believe in either statues or lumps—also object to this story. For they deny that *two numerically distinct* physical objects could be 'wholly co-located'. That is, they deny that two distinct physical objects could be composed of exactly the same parts at some level of decomposition. The lengths to which philosophers have gone to avoid co-location—from endorsing relative identity to embracing mereological essentialism—testify to just how objectionable they find it.[19]

[19] Geach (1980) rejects co-location on the way to relative identity. Van Cleve (1986) and Zimmerman (1995) reject it on their way to mereological essentialism. Others who reject co-location, on the way to their own striking conclusions, include Sidelle (1998), van Inwagen (1981), Heller (1990), Burke (1992, 1994), and Rea (2000).

Co-location is a *mereological relation*. It is the sharing by two numerically distinct objects of all their parts at some level of decomposition. It is not—as the name might misleadingly suggest—a spatial relation.[20] One can oppose mereological co-location while happily accepting, for example, that an organism is located exactly where a region of space or an event is.

Co-location does not require the sharing of all parts, but only of all parts at some level or other of decomposition. So the statue and the lump would be co-located if they were composed of, for example, exactly the same atoms. Such colocation is consistent with the statue's having a proper part that the lump lacks. For instance, the statue might have a 'hand' as a part while the lump's corresponding part might be a numerically distinct hand-shaped lump.

To see why some resist co-location, consider a dog and its atom-for-atom duplicate. Presumably, that the microstructure of one dog is qualitatively identical to the microstructure of the other implies that one could survive some misadventure, such as being squashed as flat as a pancake, if and only if the other could. Similarly, some philosophers object to colocation of the statue and the lump since it seems like their qualitatively (because numerically) identical microstructure should rule out their having different persistence conditions.[21] And because co-located objects are invariably supposed to

[20] There is no less misleading name available. The primary alternative to 'co-location' is 'coincidence'. Yet the relevant ordinary meaning of 'to coincide' is to occupy the same space at the same time. 'Constitution' is no help either, since it is not usually used as a synonym of 'co-location'. To see this, note that colocationists usually take constitution—unlike co-location—to be asymmetric: the lump allegedly constitutes the statue, not vice versa (see Baker 2000, ch. 2; Doepke 1982).

[21] The claim that the lump could survive being squashed, while the statue could not, is in one way *more* striking than the claim that the duplicate dog, but not the original, could survive squashing. For the lump and the statue have not only the same microstructure, but also—unlike dog and duplicate—all (or virtually all) the same *relational* properties as well. (See Zimmerman 1995: 87–8; Burke 1992; and Sosa 1987, §G.)

differ in persistence conditions, we have here a reason to object to co-location itself.

(This objection to co-location raises special problems for the many *David*-candidates (§II). For the *David*-candidates, just like our duplicate dogs, are objects *of the same kind*. And even some avowed co-locationists—such as Locke (1975: 328) and Wiggins (1968)—deny the possibility of co-located *kindmates* with different persistence conditions. On the other hand, belief in co-located objects of differing kinds—unlike belief in the *David*-candidates—commits one to the claim that those objects' kind-membership supervenes on neither microstructure nor environment nor anything else they share.)

There is another worry with the claim that a statue is co-located with a numerically distinct lump. This claim seems to imply—as far as causal explanations are concerned—a *needless multiplication* of physical objects. For the lump, once we have the statue, seems to bring no new causal powers into the world. Likewise, it seems that everything the alleged statue causes would already be accounted for by the work of the statue-shaped lump, which everyone treats as if it were a statue (*pace* Baker 2000: 20–1).

The aversion to differences in persistence conditions among microstructural duplicates, along with a distaste for 'needless multiplication', motivates the rejection of co-location. Moreover, near the close of Chapter 3 I'll give a new argument against co-location. At any rate, I shall proceed on the assumption that numerically distinct physical objects cannot be wholly co-located.[22]

Given that assumption, whenever there is an alleged statue and an allegedly co-located lump of clay, at least one of the following is true:

[22] Of course, there are philosophers who embrace co-location and are unconvinced by considerations raised against it. These include, in addition to Locke and Wiggins, Lowe (1983), Baker (1997, 2000), and Johnston (1992). More extensive lists of co-locationists can be found in Burke (1992, n. 1) and Baker (1997, n. 3).

(1) There is no statue.
(2) There is no lump.
(3) The lump (exists and) is *identical with* the statue.
(4) Both statue and lump exist, but they are *not* wholly co-located (i.e. they do not have all the same parts at any level of decomposition).[23]

Eliminativism implies (1) and (2), thus providing a double dose of protection from co-located statues and lumps. And I will argue that eliminativism, even with its inelegant overkill, is the best response to the fact that at least one of (1) to (4) is true.[24]

Another response to this fact is to claim (1) is true, but not (2). This response seems objectionably *arbitrary*—why deny the statue exists but not the lump? Or suppose one denies the existence of the lump but not of the statue (i.e. endorses (2) but not (1)). This too seems arbitrary.

To better understand this charge of arbitrariness, and to better see the metaphysical profligacy of co-location, consider the following. The aggregate of atoms, so the co-locationist's way of thinking suggests, cannot survive the loss of a single atom. So its persistence conditions differ from that of the statue and also that of the lump. (The lump allegedly survives the loss of an atom but not the loss of any piece of clay.) So the aggregate is numerically distinct from both the statue and the lump. Thus the sort of reasoning that leads to co-location of lump and statue leads to co-location of lump, statue, and aggregate of atoms. Presumably, this sort of reasoning leads to the co-location of more than three objects. (Van Inwagen 1990: 126 argues that those who believe that statues are co-located with

[23] One of these four is true given classical assumptions about identity. But those who adopt, for example, the 'Aristotelian' view suggested, but not endorsed, by Rea (1998*b*) might deny all four.

[24] As Michael Rea (1995) has shown, any response to the problem of the statue and the clay will generate a general response to what he calls 'the problem of material constitution' and sundry other puzzles, including the Ship of Theseus. For the record, my response to the statue and the clay denies the conjunction of Rea's Existence Assumption and his Essentialist Assumption.

lumps should believe that an *infinite* number of objects are thus co-located.)

Given our rejection of co-location, those who believe in statues must deny the existence of *any* of the objects allegedly co-located with statues. But, for example, it is objectionably arbitrary to insist, without argument, that atoms arranged statuewise compose a statue but atoms arranged lumpwise do not compose a lump and atoms arranged aggregatewise do not compose an aggregate. Similar problems afflict believers in lumps.

So far I have focused on avoiding co-location by affirming one or both of (1) and (2). Yet another way to avoid co-location is to endorse (3), the claim that the statue is *identical with* the lump. Suppose this claim is true. And suppose that sometime in the future the statue (= lump) is squashed. Does that statue (= lump) survive? To say that it does indeed survive is to open oneself up to the charge of arbitrariness. For this is, in effect, to say that that object has the persistence conditions of a lump (but not a statue). And at first glance, there seems to be nothing more than a mere verbal disagreement between one who says that only the lump, not the statue, exists and another who says that the lump is the statue, but behaves exactly like a lump and not at all like a statue. (Similar remarks apply, of course, to one who says that the statue (= lump) behaves like a statue.)

But some will tell us to take a second look. They will say that *being a statue* is a 'phase sortal' of the lump just like, for example, *being a child* is a phase sortal of a human. A child is identical with a human—there is just one person there—but ceases to be a child as the human ages. Likewise, they will say, a statue is identical with a lump—there is just one object there—but ceases to be a statue when the lump is squashed.

Suppose *being a statue* is a phase sortal. Then there is more than a verbal disagreement between the claim that the statue (= lump) survives squashing (but ceases to be a statue) and the claim that the statue does not exist in the first place. For those

who endorse (3) on the grounds that *being a statue* is a phase sortal no more deny the existence of the statue than do we who think *being a child* is a phase sortal deny the existence of children.

Allegedly, statues can survive changes—such as the loss of a bit of clay—that lumps cannot. But nothing parallel is alleged about children and humans. That is, no one thinks a child can survive changes a human cannot. So it is not obvious that the cases are relevantly analogous. Moreover, there had better be a good reason to conclude that *being a statue,* rather than *being a lump,* is a phase sortal. For if there is not, our imagined objectors fall prey to a charge of arbitrariness. (And what do they say about *being an aggregate of atoms*?)

Now I concede that many who reject co-location in the context of their own ontology of material objects have adequately responded to the charge of arbitrariness. For example, armed with a principled argument for mereological essentialism, one can non-arbitrarily avoid co-location by denying the existence of every alleged object (including the statue) other than the one supposed to have all of its parts essentially. So there are strategies other than mine for avoiding co-location in a non-arbitrary manner. But as none of these has been universally endorsed, I want to present the eliminativist's way of doing so. And I hope that in the context of the overall defence of eliminativism presented in this book, that way—which is not open even to a *prima facie* charge of arbitrariness regarding the statue and the lump—will seem to some (as it does to me) to be the most attractive.

Finally, consider (4), the claim that the statue and the lump both exist yet are not wholly co-located.[25] To understand (4),

[25] Those who endorse (4) do *not* claim that, at a certain time, the statue and lump share all their parts except, for example, a single atom. Such a claim has troubling co-location-like implications, implying that two objects which differ only trivially in microstructure—and in a way (like an extra atom atop the statue's head) that doesn't generate a macroscopic difference—have substantially different persistence conditions (cf. n. 1 above).

recall from the previous chapter that the perdurantist (a.k.a. four-dimensionalist) denies that persisting objects are wholly present at any one time. She believes, instead, that such objects exist at a time by having a proper (temporal) part that is wholly present at—that is, that exists in its entirety at—that time. Thus the perdurantist who believes a lump of clay exists on both Monday and Tuesday would say that that lump has a part that exists on and only on Monday and a distinct part that exists on and only on Tuesday. The lump itself, according to the perdurantist, never exists in its entirety on any single day. Rather, she says, it is spread out over time much as a spatially extended object is spread out over space, never wholly present at any one point.

Consider the lump's Monday temporal part. The lump's Monday temporal part is a mere *proper* part of that lump. Suppose further that a statue has that very same temporal part as *its* Monday temporal part. Finally, suppose that the lump has many other temporal parts, such as its post-squashing temporal parts, that it does *not* share with the statue. Given all these suppositions, it follows that, although the statue and clay are co-located *on Monday*—they share their Monday temporal part (and its parts)—they do not share all their parts *simpliciter* and so are not *wholly* co-located.

Given that the statue and the clay share only some parts, the claim that both of them exist and are numerically distinct is not subject to the charge of needless multiplication. And given that they have differing parts, there is nothing obviously objectionable about the claim that they have different persistence conditions. Thus the problems with the thesis that distinct objects are *wholly* co-located do not seem to afflict the claim that distinct perduring objects are co-located *at a time*.

So perdurance initially seems to provide an elegant way of avoiding objectionable co-location while at the same time allowing that a statue and a lump may be temporarily co-located. But, as others have noted (e.g. Gibbard 1975), perdurance

cannot, all by itself, block every threat of objectionable co-location. Imagine two lumps fashioned into 'statue-halves' and then stuck together to create a statue and also a large lump. Imagine also that the resultant statue and large lump are then simultaneously annihilated. In such a case it seems the statue and the lump are *wholly* co-located, sharing *all* their parts (at some level of decomposition), temporal or otherwise.

The perdurantist could respond that, if the statue is *wholly* co-located with the lump—if they share *all* their temporal parts—then the statue is *identical* with the lump; otherwise not. So in such cases she recommends response (3). But, even given a perdurance ontology, (3) faces a version of the original charge against it, that of arbitrariness. For consider the alleged four-dimensional statue (= lump). *Could* it survive being squashed? Does it have the property of *possibly surviving being squashed*?

Suppose it does. Then the statue (= lump) has the modal properties normally associated with a lump, but not with a statue. I don't see any substantive difference between this claim and the claim that there is a (statue-shaped) lump but no statue. And so the charge of arbitrariness returns. Indeed, that charge is compounded in this context, given that the perdurantist in question thinks that there are both statues and lumps. So it cannot be a problem with statues *as such* that justifies her denying a statue's existence in a case of apparent complete temporal overlap with a lump. (Similar points apply, of course, if we say that the statue (= lump) could not survive being squashed.)

Some might object that saying that the statue (= lump) could survive being squashed is not tantamount to denying the statue's existence. It is instead, they claim, merely to deny the statue is a statue essentially. Even if they are right, a worry about arbitrariness remains. For it still seems at least prima facie arbitrary to say that the statue (= lump) is only contingently a statue but essentially a lump. (A similar point holds if we say it is contingently a lump and essentially a statue.)

The perdurantist who believes in statues and lumps can avoid all charges of arbitrariness only by claiming that modal predicates are *inconstant,* expressing different properties in different contexts (Lewis 1986a: 248 ff.). The clearest account of such inconstancy comes from sortal-relative counterpart theory, so let us focus on that.

The counterpart theorist would say that there is one four-dimensional object in question, the statue (= lump); that it has a counterpart *qua* lump which survives being squashed; and that it has a distinct counterpart *qua* statue that does not. And in this way she can say that the statue (= lump) can—considered *qua* lump—survive being squashed but also cannot—considered *qua* statue—survive being squashed. This gets us inconstancy. In contexts in which we think of the object as a statue, the predicate 'possibly survives being squashed' expresses the property of standing in the statue-counterpart relation to a squashed survivor. In contexts where we think of it as a lump, the predicate expresses the property of standing in the lump–counterpart relation to a squashed survivor.

Consider the perdurantist's spin on the *David*-candidates. Many four-dimensional candidates share the temporal parts at play in the early stages of the Game. Some survive the annihilation of atom A and others do not because some have post-annihilation-of-A temporal parts and others do not. But suppose *David* had been annihilated right at the start of the Game, so that 'the candidates' are identical; that is, suppose there is but one candidate. Could it survive the annihilation of atom A? The perdurantist's reply will presumably include the claim that 'possibly surviving the annihilation of atom A' is inconstant.

There are some ordinary contexts in which we think of a statue (if statues exist) *as a statue* and others in which we think of it *as a lump*. These contexts make the perdurantist-cum-inconstant-modal-predicates solution to the puzzles involving supposed co-location more plausible than it would otherwise

be. Yet there do not seem to be the requisite ordinary contexts for *David* in the case where—due to instant annihilation—there is only one *David*-candidate. There do not seem to be ordinary contexts in which we think of it as candidate$_1$, others in which we think of it as candidate$_2$, and so on for *all* the requisite candidates. And without these contexts, I don't think one can make good on the claim that *David*'s modal predicates are inconstant in the relevant ways. So I think we should reject the 'perdurance-plus-inconstant-modal-predicates' spin on the alleged *David*-candidates.

More generally, many of us reject either perdurance or the inconstancy of modal predicates—or both. For us, I think reflections on the statue and the clay, and on co-location in general, support eliminativism.

I shall defend the existence of human organisms. I avoid the co-location of a human organism and a lump of tissue by eliminating lumps of tissue. Do I not then face the charge of arbitrariness? Perhaps, but only prima facie. And, in Chapter 5 (§III), I answer that charge by presenting reasons to believe in human organisms. These reasons, we shall see, are not also reasons to believe in lumps of tissue, or in any other object alleged to be numerically distinct from, but co-located with, human organisms. (In that same section I also discuss atoms versus 'lumps of atom-stuff'.)

IV. Brains and Thinkers

John Locke claimed that persons were one thing, thinking substances another. So, according to Locke, even if each human person had a soul of the sort substance dualists like Plato, Augustine, and Descartes believed in, a person would not be *identical* with his or her soul. For persons, Locke argued, persist by way of continued consciousness, whereas

thinking substances do not. Indeed, Locke thought it possible that one and the same person could be associated with various thinking substances at various times during his or her life and, conversely, that one thinking substance could, at different times in its career, be associated with different persons.

Nevertheless, Locke thought that the person was intimately related to her thinking substance. He thought a thinking substance had the same mental properties as 'its' person (Locke 1975: 328 ff.). Indeed, these are the *only* properties—save its brute persistence conditions—a thinking substance is alleged to have. Stripped of those mental properties, it would be a completely unintelligible I-know-not-what.

Locke's thesis implies that, whenever I have a thought, *two* thinkers have that thought: me and 'my' thinking substance. But it seems wrong that where we normally think there is one thinker—one conscious, reflective entity—there are really two. In part for this reason, Locke's distinction between thinking substance and non-physical person has proven unpopular among substance dualists. I can't think of a single one who embraces it.[26]

Dualists have rejected Locke's distinction between non-physical person and thinking substance and its resultant multiplication of thinkers. They can do so at little or no cost. But, perhaps surprisingly, materialists can reject a similar distinction only by departing in one way or another from folk ontology.[27] For if *brains* existed, they would play a role akin to that of Locke's non-personal thinking substance. With this in

[26] But a view with the same difficulties is accepted by some dualists (e.g. Swinburne 1986: 146), the view that a person is a 'compound' of soul and body. If the soul has the same mental properties as the 'compound' person, then we have twice as many thinkers as persons. Moreover, this view implies that I cannot tell whether I am a soul or a compound; after all, things seem exactly the same to both; thus I can't tell whether or not I am a person or even whether I am spatially extended.

[27] And only by departing from the standard perdurantist claim that I think that *p* at a time if and only if I have a temporal part that, at that same time, thinks that *p* (cf. Ch. 4, §III).

mind, consider the following four claims which seem to be implied by materialist folk ontology:

(1) Within the region filled by atoms arranged (normal, healthy, awake) human organismwise, there is exactly one conscious entity.
(2) Any object with atoms arranged (normal, healthy, awake, human) brainwise among its proper parts is conscious.
(3) Within the region filled by atoms arranged human organismwise, there is a human organism that has atoms arranged brainwise among its proper parts.
(4) Within the region filled by atoms arranged human organismwise, there is a brain that has atoms arranged brainwise among its proper parts.

If (2), (3), and (4) are true, then within the region occupied by a human organism there is a conscious human organism and a conscious brain. The brain is not identical with the organism; they differ in properties and parts. So (2), (3), and (4) imply that there are at least *two* conscious entities within that region. And so they imply that (1) is false. Thus at least one of (1), (2), (3), or (4) is false. (For similar reasoning, see Olson 1995: 187.) The moral I draw is that (4) is false. I have atoms arranged brainwise. But I do not have, in addition to those atoms, a brain. (Reader inserts wisecrack.)

Let us consider the other options. Suppose (2) to (4) are true and (1) the culprit. This commits one to the worst of the Lockean thesis. It implies that my thoughts are not mine alone, but shared, quite literally, by a three-pound object inside my skull. And so it is for all of us. And thus we have the unacceptable multiplication of thinkers.

This multiplication of thinkers leads to a further problem. Suppose (2) to (4) are true. Suppose both my brain and I think. Suppose my brain and I think the *same thoughts*. But this implies that I can't tell whether I am an organism or a brain.

After all, when I confidently think to myself 'I am not a mere brain but instead a human organism', my brain thinks the same thing to itself, with equal confidence. And so on for all that I think and believe and feel and experience. For all I can tell, then, I might actually be a brain. No mere sceptical hypothesis, the odds of each thinker's being a brain in a (cranial) vat, on the view under consideration, are fifty-fifty! (We could likewise object that, on Locke's view, one cannot tell whether one is a person or a mere thinking substance.)

One might object that 'I' can only have a person as its referent (see Noonan 1989: 75–6; Heller 2000). A brain's 'I-thoughts', so one might insist, refer to its 'associated' person. So whenever one has the thought 'I am a human organism, not a mere brain', then—given that human persons are human organisms—one can be confident that that thought is true.

Just for the sake of argument, let us concede this objection. This objection makes things worse. For it implies that you cannot tell whether your 'I-thoughts' refer to you or, instead, to the person in whom you are encased. And the referent of 'I-thoughts' aside, the point remains that if the brain and the human have all the same mental states, then they have all the same *phenomenology*. Things *seem to* the brain just as they *seem to* the human. That alone implies that one cannot tell whether one is a human being or a brain.

I could conclude that I can't tell whether or not I am a brain. And I could conclude that wherever we normally think there is one thinking conscious entity, there are two. These conclusions are less happy, I say, than the conclusion that although I have atoms arranged brainwise, they fail to compose a further object. So it is better to deny (4) than to deny (1).

Let us turn to (2). (2) says only that objects with atoms arranged brainwise as parts are *in fact* conscious. (So (2) is not threatened by the metaphysical possibility of 'zombies'; see Chalmers 1996: 94–9.) Denying (2) might lead to an unwarranted scepticism about who, or what, is conscious. Moreover,

those who reject (2) (yet accept (1), (3), and (4)) should offer a replacement that would explain why having atoms arranged brainwise, for some kinds of things (like me) is connected to consciousness but not so for other kinds of things (like my brain). So I think we should accept (2).

Some might object that *merely* having atoms arranged brainwise among an object's proper parts is not sufficient for that object's being conscious. Rather, they might say, those atoms must be 'appropriately integrated' among the object's other parts. The unrestricted compositionist might illustrate this point by saying that my atoms arranged brainwise are not appropriately integrated among the parts of the object composed of me and the Eiffel Tower. Thus, she says, I am conscious but not that object.[28]

Suppose we concede this objection about appropriate integration. Qualify (2) in the relevant way. Still, on any plausible qualification, the atoms arranged brainwise would be appropriately integrated among the parts of my brain—if my brain existed. So, even the relevantly qualified (2) is inconsistent with (1), (3), and (4). So, assuming the relevantly qualified (2) is true, we still must reject (1), (3), or (4).[29]

Moreover, believers in brains—even those who want to quibble with how (2) is stated—presumably want to say that brains are the things in us that think, just as hearts are the things in us that pump blood. Thinking is a large part of what brains, if there are any, are supposed to *do*. So, if there is a brain in me, then there is a thinker in me. If that brain is not identical with me—if I weigh more than three pounds—then

[28] I draw a different moral: I conclude that if such an object existed, I couldn't tell whether I were that object with the Eiffel Tower as a part or instead a human. But I can tell. So there is no such object. But never mind.

[29] One might object to (2) (recast in terms of appropriate integration) on the grounds that our concept of consciousness is 'maximal', meaning roughly that it applies only to objects that are not themselves proper parts of other conscious objects. I discuss this claim in Ch. 4 (§III). It should be clear from what I say in Ch. 4 how I would defend the argument of this section against that objection.

there are two thinking things within the region I occupy. So again, even if one doubts (2) as worded above, one should still deny the existence of the brain—unless, that is, one is willing to deny (1) or (3).

We should seek the least extreme response to the mutual inconsistency of (1) to (4). And so we are forced to deny one of the following:

(3) Within the region filled by atoms arranged human-organismwise, there is a human organism that has atoms arranged brainwise among its proper parts.
(4) Within the region filled by atoms arranged human-organismwise, there is a brain that has atoms arranged brainwise among its proper parts.

For now, for the purposes of motivating eliminativism and undermining folk ontology, seeing that either (3) or (4) is false is good enough. For either way, we have a striking example of things arranged F-wise but no Fs. My defence of the positive part of my ontology, the claim that there are human organisms, comes later. And in Chapter 5 (§IV) I'll explicitly argue that we should deny (4) rather than (3).

We can offer one further argument in support of eliminating brains. Suppose my 'brain' is put into a new 'body'. I believe that I go with my 'brain'. Yet I also believe that I am an organism, not a brain. I think the best way to render these beliefs consistent is to interpret a case of 'brain transplant' in the following way (cf. van Inwagen 1990: 169–81). I am 'whittled down' to brain size. This is the logical extreme of amputation and the resultant shrinking of the amputee. When put in a 'new body', I grow rapidly as new parts are added to me. This is the logical extreme of receiving transplants and the resultant growth of the recipient.

So suppose I am whittled down to brain size. What, then, is my relation to the brain that—before whittling—was a proper part of me? Any answer one might give—e.g. the brain and I

are co-located or my shrinking causes the brain to disappear—is problematic. Perhaps it is best to deny the existence of the brain in the first place. The puzzle thus dissipates (cf. van Inwagen 1981).

Consider one last puzzle involving human organisms, the 'problem of the corpse'. Suppose I shall cease to exist at death. (This supposition is controversial, rejected by, for example, Carter 1999.) Folk ontology says a corpse, an organism-sized material object, will lie in my coffin. But if I shall not exist, I shall not exist as a corpse. Whence the corpse? The response that the death of a human organism generates a new, human-sized, physical object is implausible (see Shoemaker 1999*a*: 500). One might instead respond (along with Sosa 1987: 156–7) that since my body will indeed be a corpse, yet I will not, I am not identical with my body; but this points towards either substance dualism or the co-location of a body and a numerically distinct person. Maybe the best response is that there are no corpses, but rather only atoms arranged corpsewise. And that's just what eliminativism implies.

V. Conclusion

One could respond to the Sorites Game by embracing metaphysical vagueness. Or one could avoid commitment to co-location by endorsing mereological essentialism and the claim that exactly one object—a mereologically invariant one—is composed of atoms arranged statuewise–lumpwise. And so on. Moreover, eliminativism itself is a striking thesis. Thus it is far from obvious, one might object, that eliminativism is a more plausible response to the cases presented in this chapter than are any of its rivals.

In partial response to this objection, I could note that if one rejects substance dualism and perdurance, but thinks persons

persist for any appreciable duration, then presumably one must reject mereological essentialism. And if one thinks metaphysical vagueness solves nothing because all the original problems reappear in the form of higher-order vagueness, one won't see in it a response to the Sorites Game. And I could argue that eliminativism handles *all* of the above cases, which no other single view does, at least no other view that doesn't have problems eliminativism avoids. And so on. And on. And on.

I have, of course, noted non-eliminativist ways to respond to the puzzles considered above. And I have raised some concerns with some of these other responses. But I shall not attempt to say everything that can be said for and against every possible ontology, even when such ontologies bear on the considerations raised in this chapter. For my primary aim has not been to demonstrate that eliminativism is far superior to any possible rival. It has rather been to show what eliminativism can do, emphasizing that its ability to do these things is a mark in its favour.

Moreover, its ability to do these things shows that eliminativism makes sense. Thus this chapter should have banished completely any residue of suspicion, not purged in the first chapter, that eliminativism is contradictory or incoherent or trivially false. For in so far as we understand the eliminativist's solutions to the puzzles suggested above—solutions requiring, for example, atoms arranged statuewise but no statues and atoms arranged brainwise but no brains—we understand eliminativism.

Eliminativism's coherence is established. Now the most significant threat to arguments for eliminativism is the reaction that, because eliminativism is *so* counter-intuitive, any such argument (if valid) should be taken to show that one of its premisses—even if they are all initially compelling—must be false. This threat stems from the overwhelming feeling of obviousness attached to the claim that statues and brains exist.

I hope that the arguments in this chapter have changed how you feel about eliminativism. I hope the arguments of this chapter have made the claim that statues exist—in addition to atoms arranged statuewise—seem somewhat less overwhelmingly obvious than it might have initially seemed. I hope I have, at the very least, weakened the conviction that eliminativism is false.

With this in mind, just try to imagine a world like ours except that, while there are atoms arranged statuewise in that world, there are no statues. Just try to imagine a world in which the correct responses to the puzzles I have considered in this chapter are the responses I have defended, responses predicated on eliminating the problematic objects. Such a world would *seem to us* just like the actual world. No amount of looking around could distinguish that imagined world from ours. But in that world the truth dissolves many philosophical puzzles: the puzzles are shown to have rested on a mistake.

Now ask yourself: is it overwhelmingly obvious that this imagined world isn't the actual one? Is it so obvious that no argument could convince you otherwise? The last two chapters have been successful if, though you still deny eliminativism, you grant that its denial is not overwhelmingly obvious, not so far beyond the pale as to invert automatically arguments for eliminativism into arguments for the falsity of some of their own premisses.

3 Epiphenomenalism and Eliminativism

Consider the following argument about an alleged baseball causing atoms arranged windowwise to scatter, or, for ease of exposition, causing 'the shattering of a window'.[30]

(1) The baseball—if it exists—is causally irrelevant to whether its constituent atoms, acting in concert, cause the shattering of the window.
(2) The shattering of the window is caused by those atoms, acting in concert.
(3) The shattering of the window is not overdetermined.

Therefore,(4)

If the baseball exists, it does not cause the shattering of the window.

The rest of this chapter will, in one way or another, involve this argument, which I shall call 'the Overdetermination Argument'. I shall begin by defending its validity, and then proceed to explicate, and defend the truth of, each of its

[30] I use 'the shattering of a window' as a plural referring expression, shorthand for many scatterings. I am not identifying the many scatterings with some single event, a shattering; that would imply that identity holds one–many. Nor do I claim that 'the shattering of a window' normally means many scatterings.

premisses. I shall conclude by arguing that the moral of the Overdetermination Argument is the truth of eliminativism.[31]

I. The Causal Principle

Suppose some individuals, such as the members of an unruly mob, cause the vandalism of a park. Suppose also that the vandalism of the park is not overdetermined. And, finally, suppose that I am 'causally irrelevant' to whether those members cause the vandalism.

This final supposition invites me to explain 'causal irrelevance'. Causal irrelevance, as I shall understand it, amounts to exactly four things. Those four things, applied to this particular case, are as follows. First, I am not myself one of the members. Second, I am not a 'partial cause' of the vandalism alongside the members; that is, it is not the case that only when combined with my additional causal contribution do the members cause the vandalism. Third, I am not an intermediate in a causal chain between the members and the vandalism; that is, the members do not cause the vandalism by causing me to do something by which I, more proximately, cause the vandalism. And, finally, I do not cause any of the members to cause the vandalism.[32]

[31] The Overdetermination Argument resembles a familiar overdetermination-based argument in the philosophy of mind, a version of which is advanced by, among others, Jaegwon Kim (e.g. 1989*b*). But there are significant differences between the Overdetermination Argument and Kim's. Those differences will become clear in Chapter 6 (§I), where I argue that—though the Overdetermination Argument is sound—Kim's is not.

[32] One may interpret this last clause as implying that I cannot prevent the members from vandalizing the park. And the first clause—my not being one of the members—can be read as an instance of a more general constraint on causal irrelevance: the x s are causally irrelevant to whether the y s have an effect only if none of the x s are any of the y s. This makes causal irrelevance symmetric.

It should be clear, given the above suppositions, that I do not cause the park to be vandalized. And that I do not cause the vandalism is the result of the following general (and so implicitly universally quantified) principle:

> *Causal Principle*. Suppose: O is an object. The *x*s are objects. O is causally irrelevant to whether the *x*s, acting in concert, cause a certain effect E (i.e. O is not one of the *x*s, O is not a partial cause of E alongside the *x*s, none of the *x*s cause O to cause E, and O does not cause any of the *x*s to cause E). The *x*s, acting in concert, do cause E. And E is not overdetermined. It follows from all this that O does not cause E.

In this principle, and in the Overdetermination Argument, *overdetermination* is understood in the most literal, straightforward, and natural sense possible. An effect is overdetermined if the following are true: that effect is caused by an object; that object is causally irrelevant to whether some other—i.e. numerically distinct—object or objects cause that effect; and the other object or objects do indeed cause that effect. Given this understanding of overdetermination, the Causal Principle is obviously and demonstrably true.

The Causal Principle is true. As noted above, the Causal Principle implies that I do not cause the park to be vandalized. More interestingly, the Causal Principle implies that the Overdetermination Argument—given how I understand overdetermination in that argument—is valid.

Some will object that I have ignored the fact that, while I stand in *no* salient relations to the mobsters, the baseball is *composed of* the atoms. Because of this fact, they will object, even if a baseball caused an effect also caused by its atoms, it would not overdetermine that effect. Or, better, they will object that although the baseball would 'overdetermine' that effect in the sense of overdetermination I explained above, there is nothing troubling about such 'overdetermination'.

This objection must be addressed. But not here. For any objection along these lines is an objection to neither the

Causal Principle nor the validity of the Overdetermination Argument. Such an objection is—as we shall see—an objection to premiss (3) of the Overdetermination Argument. And so I shall delay responding to that objection until I defend that premiss. With this point clarified, there should be no doubt that, as I intend them to be understood, the Causal Principle is true and the Overdetermination Argument valid.

II. Atomic Causation

(1) The baseball—if it exists—is causally irrelevant to whether its constituent atoms, acting in concert, cause the shattering of the window.

Suppose that the atoms working in concert only 'partially' caused the window's shattering. Suppose further that the 'full' cause included, alongside and in addition to the work of the atoms, the work of the baseball itself. Then (1) would be false. But the baseball and the atoms are not—according to anyone—relevantly analogous to two rocks jointly shattering the window, either one of which alone could not do so. For while two rocks can do more work than one, a baseball and its constituent atoms cannot do any more than those atoms all by themselves.[33]

Suppose the atoms arranged baseballwise caused the shattering of the window by causing the ball to shatter the window. Then (1) would be false. But there is not a causal chain, starting with the atoms working in concert and ending with the shattering, which includes as an intermediary the work of the baseball. Even if the baseball caused the shattering, its doing so

[33] (1) does not imply that the baseball is not a partial cause of the window shattering. (That would presuppose (4).) (1) implies only that the baseball is not a partial cause alongside and augmenting *its constituent atoms*.

would not be akin to its being the middle domino in a row of three, the domino whose falling allows the first to cause the last to fall.

Or so I say. And I think almost everyone would agree. But suppose that someone claimed that the way the baseball's atoms were at an earlier time caused the baseball, at a later time, to shatter the window. This stays within the pale just so long as he adds that the way the atoms were at the earlier time *also* caused the baseball's atoms to shatter the window at the later time.

Such an argument implies that the baseball, with respect to the shattering of the window, is causally redundant; it merely overdetermines the work of its atoms. And any (even somewhat) plausible way of defending the 'middle domino' objection to (1) requires this sort of systematic overdetermination. As I argue in defence of (3), we should resist just this kind of systematic causal overdetermination. If that argument is sound, then this objection to (1) can be blocked. On the other hand, if what I say in defense of (3) is mistaken and (3) turns out to be false, the Overdetermination Argument is sunk anyway and so this objection doesn't matter. So I shall ignore this objection in what follows.

The final point in defence of (1) is that the baseball does not cause the 'actions' of any of the atoms arranged ballwise. This rejection of 'downward' causation is part of the 'scientific attitude' and 'bottom-up' metaphysics, according to which the final and complete causal stories will involve only the entities over which physics quantifies. Of course, it is controversial whether everything conforms to a bottom-up metaphysics—I'll deny that humans do—but I think few would resist taking the 'scientific attitude' towards, and applying a bottom-up metaphysics to, baseballs (if baseballs exist in the first place).

One might object that while some sorts of 'downward causation' are forbidden by a bottom-up metaphysics of baseballs, not all are. Specifically, one might say, such a metaphysics

permits the way a baseball is at one time to cause its atoms to do something at a later time—and so permits downward causation—just as long as those atoms' doing that something at the later time also has a complete causal explanation *wholly in terms of the microphysical*. (Otherwise, as will become clear in the defense of premise (2), the baseball would have 'emergent' causal powers.) Thus—and this is essentially the same sort of move we considered in defence of the 'middle domino' objection—one might claim that bottom-up metaphysics allows downward causation if and only if downward causation *merely overdetermines* microphysical causation, if and only if it is merely redundant.

The only sort of downward causation even arguably consistent with the bottom-up metaphysics of baseballs—and so the only sort that can plausibly generate an objection to premiss (1)—implies systematic causal overdetermination. As noted above, in response to the 'middle domino' objection, I shall argue against systematic overdetermination below. For now, I shall assume that no such overdetermination occurs and ignore any objections, including the one just raised, that require it.

Baseballs do not exercise downward causation upon their atoms. Nor is a baseball an intermediary in a causal chain, bridging the work of its atoms to the shattering of the window. Nor is a baseball a partial cause, alongside its atoms, of the window's shattering. These three points above are what it is (given, obviously, that the baseball is not itself one of the atoms) for the baseball to be causally irrelevant to whether its atoms shatter the window. Given my understanding of causal irrelevance, we should be able to see that (1) is true, even uncontroversially so.

(2) The shattering of the window is caused by those atoms acting in concert.

Premise (2) seems obviously correct. After all, each of the window-striking atoms causes *something*. And when you put

what one atom causes together with what another causes, and so on for each of the atoms arranged baseballwise, it seems like the cumulative effect must be the shattering of the window.

Suppose someone denied (2) and claimed that our imagined shattering is caused, not by atoms, but by the baseball. He must claim, then, that the baseball causes something that its parts, working in concert, do not. I suppose his idea must be that the baseball causes things in virtue of having some sort of causally efficacious 'emergent' property.[34]

C. D. Broad questioned the assumption of 'Mechanism', the assumption that *every* composite object is causally redundant because it is related to its parts as a clock is to its 'springs, wheels, pendulum, etc.' (1925: 60). But we can oppose the above objection just so long as, were there clocks (or baseballs), they'd be related to some of their parts as a clock is related to its springs, wheels, pendulum, etc. That is, we can oppose the above objection just so long as truly non-redundant causal properties—properties that would allow an object to cause what its parts do not—do not 'emerge' *at the level of artefacts*.[35] And so even opponents of full-blown Mechanism (or full-blown bottom-up metaphysics) should oppose the above objection. For even they should agree that everything a baseball causes is caused by its parts at some level of decomposition.

Any objection to (2) that insists that the baseball—but not its parts—causes the window to shatter is mistaken. For any such objection implies the false claim that baseballs have 'emergent' causal powers. Nevertheless, I'll respond to two more objections

[34] If we deny (2), we open up the possibility that the baseball, in virtue of its 'emergent' causal properties, causes its atoms to do its bidding. Thus denying (2) might undermine (1). (More on the connection between 'emergent' or 'non-redundant' causal properties and downward causation in Ch. 4 §V.)

[35] So we need not rule out all non-redundant or emergent causal properties, not even all such purely physical properties. So nothing I defend here is threatened by the apparent evidence Teller (1989) and Maudlin (1994: 210–12) discuss for something like emergent causal properties in physical systems.

of this sort. We have just seen one reason these objections fail. I'll point out further problems with them below.

The Overdetermination Argument shows that *if* the atoms shattered the window, *then* the baseball did not. So, one might object, since baseballs shatter windows, their constituent atoms do not. In response, no matter what we decide to say about the window's shattering, there will be some things that the atoms seem to cause for which the baseball cannot account. Imagine, for example, the causal effects of the atoms before, or after, they (allegedly) compose the baseball (or, for that matter, when they compose nothing at all). But the converse does not hold. That is, everything that is allegedly caused by a baseball can be accounted for by the work of the atoms that compose it at various times. This asymmetry gives us strong reason, when forced to choose, to favour the causal powers of the atoms over those of the baseball.

Suppose one tried to resist this asymmetry. Suppose one argued that whatever seems to be caused by the parts of the baseball is instead—somehow—caused by the baseball itself. In § IV I'll argue that 'epiphenomenal' material objects ought to be eliminated. Thus rendering the baseball's alleged parts causally inefficacious implies that the baseball is a simple. The claim that atoms arranged baseballwise fail to compose a baseball might be hard to swallow. But it goes down like draught Guinness compared to the claim that baseballs are simples.

Here is a second objection to premiss (2). Just as the baseball is not identical with its constituent atoms, so the shattering of the *window* is not identical with the many scatterings of the *atoms arranged windowwise*. Suppose, then, that the scatterings of the atoms are caused by the atoms arranged ballwise but the shattering of the window is caused by the baseball. This would give us distinct effects with distinct causes, allowing the ball to shatter and the atoms to scatter and neither to overdetermine the work of the other.

There are two important things to note about this objection. First, this objection requires a shift in terminology. Up to this point I have been using 'the shattering of the window' as a plural referring expression, referring to the many scatterings. This objection requires that we now use it as a name of a single, composite event. So let us do so for the remainder of this section.

Second, the 'non-identity' crucial to this objection is *not* between the (single) event of the scattering of the atoms and the (single) event of the shattering of the window. The first event—if it existed—would be identical with the second (cf. Kim 1998: 83–7). For 'the scattering of atoms' would be just another description, a 'microdescription', of the shattering of the window.

There is, however, a lack of identity between the shattering of the window and the many events such as this atom's heading thataway and that atom's heading thisaway, and so on, for each of the atoms formerly arranged windowwise. Let us refer to those many events collectively as the 'multiple scatterings', being careful to remember that 'multiple scatterings' refers to many events, not to a single event composed of those many events. Given our rejection of composition as identity (Ch. 1, §IV), the multiple scatterings cannot be identical with the shattering of the window that they allegedly compose. They are many; it—assuming there is such a composite event—is one.

My response to this second objection to (2) begins by noting that the atoms arranged baseballwise have *multiple* effects, namely the multiple scatterings. And so the most reasonable thing to say is that the atoms' multiple effects include, in addition to the multiple scatterings, the shattering of the window. Moreover, the multiple scatterings *compose* the window's shattering. With this in mind, consider the following:

If some objects cause events $v_1 \ldots v_n$, and $v_1 \ldots v_n$ compose event V, then those objects cause V.

This principle, which I think is correct, implies that if the atoms cause the multiple scatterings, and if there is a composite event of the window's shattering, then the atoms cause that composite event.[36]

This completes my discussion of objections to premiss (2) that require the baseball to cause things its parts do not. But there is a very different sort of objection to premiss (2)—the premiss claiming that the shattering of the window is caused by the atoms arranged baseballwise—worth addressing. This is the objection that only events, not atoms or any other objects, cause things to happen. Now this 'objection' is most plausibly interpreted as merely reminding us that the sense in which events cause things differs from, and is perhaps more basic than, the sense in which objects cause things. But this reminder has no adverse implications for (2). Indeed, a fundamental distinction between event-causation and objectcausation strengthens the arguments of this chapter (see §III).

Of course, we could interpret this objection as the claim that there is *no sense* in which objects cause things. This claim implies that (2) is flat-out false. (It also implies that the conclusion of the Overdetermination Argument is flat-out true.) But this claim is mistaken. Consider that however the details may vary, virtually all accounts of perception agree that an object can be perceived only if it causes something.[37] Moreover, recall the familiar charge that, because abstracta would not have causal powers, they simply do not exist. Those who endorse this charge might be inclined to defend Samuel Alexander's principle, called 'Alexander's dictum' by Jaegwon Kim (1993*a*): to be real is to have causal powers.

[36] Compare Kim (1998: 42–3): '*To cause a supervenient property to be instantiated, you must cause its base property (or one of its base properties) to be instantiated.* To relieve a headache, you take aspirin: that is, you causally intervene in the brain process on which the headache supervenes.'

[37] This is true not only of causal theories of perception of the sort endorsed by Chisholm and Grice, but also of theories like Goldman's that allow non-causal factors a prominent role. (See Alston 1990 for discussion of these views.)

The point here is not to defend some particular account of perception, Alexander's dictum, or attacks on abstract objects rooted in their alleged causal inertness. The point is that the presence of these views on the philosophical playing field is ample evidence that philosophers generally—and correctly—assume that entities other than events, such as objects, cause things.

Now one might, in light of all the reasons above, accept the truth of (2), yet still be bothered by it. For one might find (2) to be an odd premiss in an argument whose conclusion is that a baseball does not cause a window to shatter. For one might think that, though a baseball is not identical with its atoms, its causing something is nevertheless analyzed as its atoms' (or other parts') causing that same thing. Thus one might object that (2) is simply *another way of saying* that the baseball causes the window to shatter.

This objection does not have the makings of an objection to the truth of (2). Nor, for that matter, will it generate an objection to the truth of (1) or the validity of the Overdetermination Argument. (It doesn't touch the validity of the Overdetermination Argument because it doesn't bring into question the Causal Principle.) So in so far as we have here an objection to the soundness of the Overdetermination Argument, it is—perhaps somewhat surprisingly—actually an objection to premiss (3). And so I shall address this objection in the next section.

III. Causal Overdetermination

(3) The shattering of the window is not overdetermined.

Consider a substance dualist (like Mills 1996) who, conceding causal closure of the physical, says that mental events cause physical events only by overdetermining the effects of physical

causes. Pre-theoretically, that's an ugly picture. The redundancy is *all by itself* a reason to resist this form of substance dualism. More generally, we always have a reason to resist systematic causal overdetermination, along with any view that implies it.

As I shall explain later, the reasoning behind the Overdetermination Argument quickly generalizes to apply to more than (alleged) baseballs and to more than the shattering of a window. Thus one who responds to the Overdetermination Argument by rejecting (3) must—assuming she wants to save the causal power of more than baseballs—embrace overdetermination in a wide variety of cases. But we should resist widespread and systematic causal overdetermination. And so I think it is most reasonable to endorse (3).

Some will disagree. Some will reply that while *certain kinds* of systematic overdetermination are surely objectionable and to be resisted, overdetermination of the sort at issue here—of the sort denied by (3)—is not. For this objection to be principled, our objector must have in mind some principled way to distinguish objectionable overdetermination from the unobjectionable. I'll consider different ways one might draw such a distinction.

To begin to understand the most plausible way of drawing such a distinction, and thus the most serious objection to (3), consider the following claim: an effect is pseudooverdetermined if it is caused by an object and caused by the event in which that object participates. This claim implies that a window's shattering is pseudo-overdetermined if it is caused by a baseball and caused by the baseball's striking the window. But, I reply, *what it is* for a baseball—if baseballs exist—to shatter a window is for it to participate in a window-shattering event. So, I say, pseudo-overdetermination is *not* overdetermination. Thus, I conclude, any objections one might have to systematic overdetermination should give rise to no objections to systematic pseudo-overdetermination.

Likewise, one might object, *what it is* for the baseball to shatter the window is for its parts—such as its atoms—to shatter the window. (This returns us to the objection raised at the very end of the last section.) So, the objection continues, the sort of 'overdetermination' opposed by premiss (3) is not real overdetermination. It is, instead, like pseudo-overdetermination. As a result, one might conclude, any scruples we have about real systematic overdetermination do not support denying the occurrence of 'overdetermination' denied by (3).

This objection assumes that the overdetermination at issue in (3) is analogous to pseudo-overdetermination. Its central assumption is that a baseball's causing something *just is* its parts causing that same thing. The first step towards seeing that this objection's central assumption is mistaken—and so towards seeing that the objection itself fails—is to consider the following:

Object O's causing an effect E *is analyzed as* O's participating in the appropriate way in an event that causes E.

Nothing turns on the details of this analysis, which is purposefully short on detail. All that matters is that in this analysis—as in *any* analysis of object-causation in terms of event-causation—*causation* appears in both analysandum and analysans. If any such analysis is to have a hope of being correct, it must not be blatantly circular. If it is not circular, then the kind of causation exercised by objects must not be the kind exercised by events. And it is quite plausible that we have different 'kinds' of causation here. To see why, it may be useful to think in terms of metaphysical categories. The causation exercised by events, since events differ categorically from objects, is only *analogically related* to the causation exercised by objects.

The above analysis can survive challenges based on circularity. For object-causation and event-causation are distinct, but interanalysed, phenomena. Again, objects and events do not do the same kind of causal work. On the other hand, a baseball

and its constituent atoms, all being objects, do the same kind of causal work. Thus an analysis of one's causing in terms of the others' causing is bound to be circular. And so any such analysis ought to be abandoned right from the start.

I suppose one might object that the analysis of a composite's causing something in terms of its parts' causing that same thing avoids circularity because 'composite-causation' and 'part-causation' are distinct kinds of causation. This objection is mistaken. As noted above, there is a significant difference between an event's causing something and an object's causing something. But there does not seem to be the same kind or degree of difference between a big object's causing something and a smaller object's causing something, even if the smaller is part of the bigger. Parts of objects and the objects they compose seem to be in the same category—object—and for that reason presumably cause things in the same sense.

And there is another problem with saying that the sense in which composites cause things is distinct from the sense in which their parts cause things. This problem stems from some composites' being themselves parts of other, bigger, composites. For the following sort of claim seems unacceptable: an atom composite-causes an effect and, when part of something else, part-causes that same effect, although these are metaphysically different kinds of causation. I suppose one might try to avoid such claims by insisting that only simples can be parts. But that too seems unacceptable. It is certainly unacceptable to anyone whose ultimate aim is to defend folk ontology, since folk ontology embraces objects with composite parts.

So here is what I conclude thus far. One can plausibly insist that *what it is* for a baseball to shatter a window is for it to participate in a window-shattering event. But one cannot plausibly insist that *what it is* for a baseball to shatter a window is for its constituent atoms to shatter it. Because an object's causing something is not analysed as its parts causing something, the

overdetermination denied by premiss (3) is *not* like pseudo-overdetermination. It is real overdetermination.

We can look at essentially the same point in this way. Because an object and an event do not do the same kind of causal work, one cannot redundantly duplicate the work of the other so as to result in overdetermination. But a(n alleged) baseball and its atoms, all being objects, do the same kind of causal work. So if they all caused the same effect—like the shattering of the window—the baseball would thereby redundantly duplicate the work of its atoms. And so the baseball and the atoms would really overdetermine that effect.

I suspect that most who were initially inclined to resist (3) were inclined to do so because the overdetermination (3) opposes seemed to be on a par with the totally innocuous pseudo-overdetermination. But now we know that the overdetermination (3) opposes is not remotely like pseudo-overdetermination. This undermines the primary opposition to (3). Indeed, it provides positive support for (3). For just as the alleged analogy between pseudo-overdetermination and overdetermination by an object and its parts supported denying (3), so the actual disanalogy speaks forcefully in (3)'s favour.

Some might still insist that systematic overdetermination by an object and its parts is not objectionable. Of course, they cannot claim to do so on the grounds that such overdetermination is not genuine. It is genuine. But perhaps they will defend the following argument. The true moral of this chapter thus far is that, like it or not, composition is possible only given overdetermination of the very sort premiss (3) denies. Since, necessarily, to be composite is to be causally redundant, this argument continues, to object to systematic overdetermination by composita is simply to object to *what it is* (in part) to be composite. But composition as such is not objectionable. Therefore, etc.

This argument is unsound. For to be composite is *not,* in its very nature, to be causally redundant, to overdetermine systematically.

As I shall argue in the next chapter, there are composite objects that are not causally redundant. It is a peculiarity of some (alleged) composita, rather than an inevitable result of composition as such, that they are wholly causally redundant. (This gives us another reason to deny that a composite's causing something is *analysed as* its parts' causing that same thing. For, obviously enough, that analysis is inconsistent with a composite's failing to be causally redundant (and thus causing something its parts do not).)

Opponents of (3) might argue in the following way. Genuine and systematic overdetermination is not objectionable if the overdetermining causes are not 'wholly separate'. The baseball is not wholly separate from its atoms. So we should not resist the claim that the baseball and its atoms overdetermine the shattering.

I think this objection fails. For I reject its claim about what kinds of systematic overdetermination are objectionable. And so do others. Consider a well-known argument in the philosophy of mind. This argument's cornerstone is that mental properties' systematically overdetermining the effects of physical properties would be objectionable (Kim 1989*a, b*; see also Malcolm 1968). Yet according to this argument, mental properties supervene on the very physical properties whose effects are in question (much as an alleged baseball supervenes on its atoms; Ch. 1, §I). The overdetermination this argument targets as objectionable would be the work of entities that are not wholly separate.

I suppose one could deny (3) on the grounds that overdetermination by a baseball and its atoms is not objectionable. This is hardly principled. Nor is there much improvement in the claim—perhaps this is what was behind the 'wholly separate' objection above—that overdetermination by an object and its proper parts is not objectionable. For this claim is tailor-made to resist a premiss like (3). And this claim loses its initial plausibility, I believe, in light of the conclusions noted above. I have

in mind here especially the conclusion that an object's causing an effect cannot be analyzed as its parts causing the same effect and the conclusion—defended in Chapter 4—that some objects cause effects without merely overdetermining what their parts cause.

We always have a reason to resist systematic and genuine overdetermination. Thus, I say, we have a good reason to endorse (3). Some might object that opposition to systematic overdetermination in general does not support (3), since there is something special about the overdetermination (3) opposes. But, I have argued, the various principled ways one might defend this objection fail; these failures rob that objection of its initial plausibility. And so we should conclude that (3) is true.

I want to add one final point in favour of (3). Imagine that someone has been killed by a bullet. Now entertain the possibility that the killing was overdetermined by two bullets arriving simultaneously. But suppose, further, that there is *no reason* to believe that the killing was overdetermined in this way. For, let us suppose, while there is evidence for the existence of one bullet, there is no evidence for the existence of a second. In such a case, I think everyone would agree that we should deny that the killing is overdetermined as a result of a second bullet. For without a reason to think an effect is overdetermined, we should assume it is not.

Obviously enough, one would have a reason for believing that the shattering of the window is overdetermined only if one had a reason for believing that *both* the baseball *and* the atoms arranged baseballwise caused it. And one would have a reason for believing *that* only if one had a reason to believe that the baseball existed. But, I shall argue momentarily, there is no good reason to believe the baseball exists. Without the positive belief that a baseball exists, there is no motivation for believing that the shattering of a window is overdetermined, caused by atoms and a ball. And if there is no such motivation, then we ought to conclude that there is no overdetermination.

And thus we have (3) of the Overdetermination Argument.

As just noted, I shall argue that we have no good reason to believe in baseballs. One might then ask why we need the Overdetermination Argument at all. Here is one reply. Even if the defence of (3) to follow fails, the Overdetermination Argument is still sound. For even if that particular defence of (3) fails, we should—for the reasons given above—still accept the truth of (3). There is a second reply. The argument to follow implies only a healthy agnosticism (not a full-blown eliminativism) about baseballs. The Overdetermination Argument shows how that agnosticism leads to the claim that baseballs, if they exist, do not shatter windows. And that, as we shall see (§IV), leads to eliminating baseballs.

Our ordinary reason for believing in baseballs is simply that, so it seems, we can just see them (or feel them or otherwise sense them). Similarly, our ordinary reason for believing in statues is that we can just see them. But we saw in Chapter 1 (§II) that 'just seeing a statue' is not really a good reason to believe that atoms arranged statuewise compose a statue. Likewise, 'just seeing a baseball' is not a good reason to believe that atoms arranged baseballwise compose a baseball. So it turns out that our ordinary reasons for believing in baseballs aren't good reasons. So unless we have some extraordinary reasons, we have no good reason at all to believe in baseballs. And if we have no good reason to believe in baseballs, then we shouldn't believe in them. (That is, we should either withhold belief or positively disbelieve in them.)

In the course of establishing, in Chapter 1 (§II), that 'just seeing a statue' isn't a good reason to believe in a statue, I traded on an analogy. I claimed that whether atoms arranged statuewise compose a statue is analogous to whether atoms arranged my-neighbour's-dogwise and the-top-half-of-the-tree-in-my-backyardwise compose an object. And I said that it would not do to support an affirmative answer to the latter question simply by saying 'I can just see that object'.

This is an important point in what follows. Note that it is *not* controversial. There are many philosophers who believe in arbitrary sums like the 'dog-and-treetop', but none of them—not one—defends the existence of such things on merely perceptual grounds. No one says we should believe that such an object exists simply because we can see it or simply because we can hear it (gnawing on a bone while rustling its leaves).

Part of the reason, presumably, that no one says such things is that one's visual and auditory experiences would be the same whether or not they were caused only by atoms arranged dog-and-treetopwise or were instead overdetermined by those atoms plus the object they compose. But whatever the explanation, it is uncontroversial that philosophical argument is necessary to justify positive belief in the dog-cum-treetop. Likewise, philosophical argument is necessary to justify positive belief in statues and, of course, baseballs.

Anyone who wants to resist this conclusion must insist that the question of whether atoms arranged baseballwise compose a baseball is *not* relevantly analogous to the question of whether atoms arranged dog-and-treetopwise compose something. Now no one should dispute that they are analogous in many ways. Each is a question about whether atoms compose a particular macrophysical object. Each is a question that, if it has an affirmative answer, has an affirmative answer of necessity (and likewise if it has a negative answer). And each of the alleged macrophysical objects, if it exists, *at best* overdetermines our sensory experience of it in *exactly the same way* as does the other.

The only possibly relevant disanalogy between the cases at issue here is that baseballs are, but 'arbitrary sums' like the dog-and-treetop are not, part of our commonsense metaphysics. In light of this, one might object that the objects of folk ontology—unlike arbitrary sums—are presumed innocent until proven guilty.

Let me concede the following. Folk ontology and belief in baseballs is a *justified starting-point* in forming beliefs about the

world. Though it is reasonable to start with such beliefs, however, their justification is undermined for those of us familiar with the issues raised in this section of this chapter. For we ought to see that the only difference between arbitrary sums and statues is a matter of conventional wisdom and local custom. *Once this is pointed out,* one is no longer justified in believing that statues exist merely because one can supposedly see them.[38]

Imagine a child reared on an island of philosophers who are enamoured of unrestricted composition. Such a child might take it for granted that arbitrary sums exist. She might even insist that such sums obviously exist because she can 'see them'. I think that child is initially justified in her beliefs. But once she realizes that she could 'see such things' whether or not they were there, seeing no longer justifies believing.

One might reply that belief in folk ontology is not merely *customary,* but somehow epistemically privileged: hard-wired in non-defective cognizers or part of epistemic proper functioning or what have you. But it's hard to see why the folk way of carving up the material world should—barring further argument—be elevated to a loftier status than the unrestricted compositionist way. Note, in particular, that the problem wouldn't be solved simply by folk-ontological beliefs' being *reliably* formed. For even if unrestricted composition were true and 'seeing arbitrary sums' reliable, such seeing would not, on its own, justify believing.

I conclude that whether atoms arranged statuewise compose a statue is in the same epistemic boat as whether atoms arranged treetopwise and neighbour's-dogwise compose an object. In the latter case, one cannot reasonably base one's conclusion simply on what one senses or, more generally, on

[38] We have here something like what Alvin Plantinga (1993: 41) calls an 'undercutting defeater' for one's non-propositional sensory evidence for the existence of a baseball. For more on defeaters, see Lehrer and Paxson (1969), Harman (1973), and Pollock (1974, 1986).

any of the alleged causal effects of the alleged treetop-plus-dog; likewise, then, in the former. To be justified in believing in baseballs, we must have philosophical reasons to believe that they exist.

Some will respond that we *do have* philosophical reasons for believing in baseballs. We can better understand their position by considering the possible worlds of David Lewis (1986*a*). Lewis believes that these worlds contain macrophysical objects, like the counterparts of our (alleged) baseball, with which we have no causal–perceptual interaction. This is germane to the present discussion because Lewis thinks he has good (non-perceptual) reasons for believing in these objects, namely, their philosophical utility combined with our modal insights.[39]

Similarly, some think that we have philosophical reasons for believing in this-worldly baseballs. Those reasons, if strong enough, would block my final defence of (3), the defence based on epistemic considerations. But once the locus of the debate moves to philosophical argument and leaves behind what we can 'just see', things look good for eliminativism. In part, this is because once we agree that the way to decide whether baseballs exist is by philosophical argument, Chapter 2's considerations against the existence of things like baseballs become all the more significant. And, in part, this is because there is very little out there by way of positive, non-question-begging arguments for the existence of baseballs. After all, their existence is generally taken for granted.[40]

[39] Some have objected that, because knowledge of concrete objects requires causal interaction, one could not know of the worlds Lewis posits or the objects in those worlds. Lewis (1986*a*, §2.4) notes, and responds to, objections along these lines found in Richards (1975), Lycan (1979), and Skyrms (1976).

[40] It is difficult to find positive arguments for the existence of baseballs. Note, for example, that Ned Markosian (1998) defends baseballs and the like only in the sense—irrelevant to present purposes—that his answer to the special composition question is *consistent with* the existence of such objects. H. Scott Hestevold (1980 –1) might get us baseball-sized and -shaped objects, but given Hestevold's avowed mereological essentialism, he does not get us *baseballs*. If sound, Crawford Elder's (1996) arguments might (perhaps) save baseballs, but only by implying that organisms and rocks, among other putative objects beloved of folk ontology, do not exist.

But there is something out there that has a right to call itself 'a non-question-begging defence of the existence of baseballs'. I have in mind philosophical defences of unrestricted composition, and specifically defences of unrestricted composition that are also committed to perdurance and the inconstancy of modal predicates.

To see the relevance of perdurance and inconstant modal predicates, consider the following. Unrestricted composition implies that, if there is a baseball B composed of atoms $A_1 \ldots A_n$, then there is some other object composed of all those atoms save A_n. Call that latter object B*. Suppose A_n then ceases to exist. What is the relation between B and B*? One answer implies that B ceases to exist. But then B cannot be a baseball. For the baseballs of folk ontology can survive the loss of a single atom. And we are here concerned with only the unrestricted compositionist who wishes to defend the existence of baseballs.

The unrestricted compositionist who defends the existence of baseballs will probably say that B and B* are co-located after A_n ceases to exist.[41] For the unrestricted compositionist who accepts co-location can easily insist that there are objects persisting in the way that the folk think baseballs do. Each of those objects, then, can be a 'baseball' with the right persistence conditions. Recall that in Chapter 2 (§III) I argued that co-location leads right to perdurance and inconstant modal predicates. So I'll assume that both are accepted by the unrestricted compositionist who embraces co-location.

[41] These comments about the implications of unrestricted composition reflect the orthodox and standard views. But they are of course controversial. To delve more deeply into these issues, see van Cleve (1986), Rea (1998*a*), and the essays in Rea (1997*b*).

The ontology of unrestricted composition and co-location includes the objects of folk ontology. But—with its explosion of macrophysical objects, massive amounts of co-location, and perdurance—it is not the ontology of the folk. So one question is whether this departure from folk ontology is more or less plausible than eliminativism. I think it is less plausible (although many will disagree). The burden of much of this book is showing just how plausible eliminativism really is, thus helping to make the case that it is more plausible than its rivals. (And I argue that it has a particular advantage over unrestricted composition plus perdurance in Chapter 7 (§III).)

If Chapter 2's arguments (augmented by those of Chapter 5) for eliminativism are only about as persuasive as the philosophical arguments for the existence of baseballs, then the philosophical arguments here end in a draw. But then eliminativism wins. For if the philosophical arguments end in a draw, we have no positive reason to believe in baseballs. And if we have no positive reason to believe in baseballs, we have no positive reason to believe that the overdetermination opposed by (3) occurs. And we should deny overdetermination occurs unless we have a positive reason to believe it does. So we should accept (3): the shattering of the window is not overdetermined.

At the start of this section I considered the objection that what a baseball causes is analysed in terms of what its parts cause. But we saw that this analysis fails. Thus we saw that the alleged overdetermination at issue in (3) is indeed genuine overdetermination. I then countered other objections to (3). And I presented reasons to accept (3). So I concluded that (3) was true, even before developing the point that, because baseballs would be at best causally redundant, none of our ordinary reasons for believing in them are any good.

At that stage of the argument—prior to my developing the 'epistemic point'—I think someone could justifiedly reply that my considerations in support of (3) are outweighed by his certainty

that baseballs both exist and cause things. After all, he might add, if baseballs exist and cause things, the Overdetermination Argument is easily transformed into an argument against premiss (3).

But we now know that the belief in baseballs is on a par with the belief in objects like the dog-and-treetop. The belief that baseballs exist (and cause things) is justified, if at all, by philosophical means. And so it merits only the degree of certainty appropriate to that of a speculative philosophical hypothesis. As a result, the denial of (3) is itself likewise speculative and thus relatively vulnerable to defeat. In light of this, I conclude that my earlier considerations in support of (3) are compelling. And so I conclude that (3) is true.

IV. The Moral of the Overdetermination Argument

I have defended the validity of the Overdetermination Argument and the truth of each of its premisses. The Overdetermination Argument is sound. Baseballs do not cause windows to shatter. And there is nothing special about shattering windows as opposed to, say, knocking hapless batters unconscious. Nor is there anything special about whether the shattering or knocking is allegedly caused by a baseball or, for example, a rock. For these reasons, the Overdetermination Argument looks like it will generalize to rob the macrophysical of causal power in a wide range of cases.

We can see how far the Overdetermination Argument generalizes by looking at the following schema of which it is an instance:

(1*) Object O—if O exists—is causally irrelevant to whether its parts $P_1 \ldots P_n$, acting in concert, cause effect E.
(2*) $P_1 \ldots P_n$ cause E.

(3*) E is not overdetermined.
Therefore,
(4*) If O exists, O does not cause E.

If an alleged effect of an (alleged) object is caused by that object's parts, and if that object is causally irrelevant to whether its parts cause that effect, then that effect is—assuming no overdetermination—not caused by the object in question. Note that, if *every* effect allegedly caused by a composite object is caused by its parts, and if that object is causally irrelevant to whether its parts cause those effects, then—assuming no overdetermination—the object causes *nothing*.

Everything (alleged) baseballs and other non-living macrophysical objects (allegedly) cause is caused by their proper parts at some level of decomposition. Moreover, if baseballs and other non-living macrophysical objects exist, they are causally irrelevant to the causing done by their atoms. At least that's what I say. For these claims are the heart of a 'bottom-up' metaphysics applied to baseballs and other inanimate macroscopica. So here is where the Overdetermination Argument and its opposition to systematic overdetermination leads us. If baseballs and other non-living macrophysical objects exist, then—since a 'bottom-up' metaphysics is true of them—they do not cause anything at all.

One is likely to wonder why I restrict this conclusion to non-living macrophysical objects. I'll gesture at my response in § V below. Developing and defending that response will be the central task of the following chapter. But this issue does not need to be settled before establishing this chapter's main thesis up to this point. For that thesis is not that humans or atoms or any other things have causal powers. It is that if non-living macrophysical objects exist, they cause nothing.

If non-living macrophysical objects exist, they cause nothing—they are *epiphenomenal*. This bears directly on eliminativism. For if there were baseballs, they would break windows,

they would injure batters, they would cause visual sensations (and so be seen), and they would cause tactile sensations (and so be felt).[42] In general, if there were inanimate macrophysical objects, they would have causal powers. But given the Overdetermination Argument and the schema of which it is an instance, if there were such objects, they would not have causal powers. So there are no such objects.

Arguments linking existence to causal powers are often controversial. Consider such arguments against Platonic Forms or moral properties. But this is no problem. For I do not rely upon the entirely unrestricted thesis that to be is to have causal powers. I claim only that, for *macrophysical objects,* to be is to have causal powers. Macrophysical objects are *exactly* the sort of things about which this kind of causal requirement seems to be true. There should be no controversy on this point. The controversy, instead, is about which other sorts of things are like macrophysical objects in this way.

We can, without too much controversy, extend this 'causal criterion' to events. (At least, we can extend this criterion to a wide range of events. Perhaps the number 7's being prime—even if that event exists—causes nothing.) And, as one might suspect, we can likewise extend the reasoning of the Overdetermination Argument to events. Imagine that my wife and I are lifting a sofa. The sofa's being lifted is a result of two distinct events: my straining at one end of the sofa and my wife's straining at the other. If, in addition to those two events, there is the single composite event of *our* straining, then that composite event would cause the sofa to be lifted only at the price of overdetermination. But that price is too high. So the composite event does not cause the sofa to be lifted. Nor does

[42] Epiphenomenalism with regard to non-living macrophysical objects implies that we have no sensory evidence for their existence. This point differs from the claim suggested during the defence of premiss (3). That claim was that we have no good reason to believe that one's sensory experience caused by atoms arranged baseballwise is overdetermined by a(n alleged) baseball. The point at hand is that a baseball does not cause one's baseballish sensory experience.

it cause anything else. But then it must not really exist. Thus my ontology of events will end up being sparse for the same reasons, and in much the same way, as my ontology of material objects.[43]

I can now redeem a promissory note issued in Chapter 2. There I said I would give a new reason to reject wholly co-located entities such as a statue and a lump of clay. Suppose, for *reductio,* that a statue and a lump of clay are numerically distinct material objects that are wholly co-located, that is, that share all of their parts at some level of decomposition. Anything the alleged statue is alleged to cause—the breaking of a window, visual sensations—would also be caused by the statue-shaped lump. Anything the alleged lump is alleged to cause would also be caused by the lump-constituted statue. But there is not the sort of systematic causal overdetermination that their co-location implies. Therefore, at least one of those objects causes nothing. But every macrophysical object causes something. So either the statue or the lump does not really exist. And so there is not co-location of a statue and a numerically distinct lump after all.

(The general strategy behind the Overdetermination Argument can, we have seen, be adapted to cases where the allegedly overdetermining competitors are an event and its 'parts' and an object and its constituting mass. The application of this sort of reasoning to a property and its supervenience base in the philosophy of mind is already familiar and will be discussed in detail in Chapter 6 (§I). There may be other areas, not explored in this book, where this strategy can be fruitfully exploited.)

[43] A sparse ontology of events supports a view I have defended elsewhere. I believe that objects endure; and this implies, I have argued (1995*a*), that events endure. That events endure is difficult to reconcile with the claim that, say, the American Civil War existed; for such an event seems never to have been 'wholly present' at any single time. But the endurance of events like my thinking that P is easier to accept.

The anti-co-locationist adaptation of the Overdetermination Argument resonates with an objection to co-location noted in Chapter 2. That objection was that co-location implies—as far as causal explanations are concerned—'a needless multiplication' of physical objects. It implies this, according to that objection, because everything one co-located object allegedly causes is accounted for by the work of the other (Chapter 2, §III). Opponents of co-location who found themselves nodding vigorously as that objection was originally presented should, in consistency, deny that atoms arranged baseballwise compose a baseball. After all, to add the baseball is to needlessly multiply. For everything the baseball allegedly does is accounted for by the work of the atoms. So at least some of what motivates denying co-location also motivates eliminativism.

V. Conclusion

I suspect that those now convinced that there are no nonliving macrophysical objects are inclined to deny that there are human organisms. For they are likely to think that the reasons to eliminate baseballs are equally reasons to eliminate humans.

The Overdetermination Argument could not be adapted to humans if humans caused things that their atoms do not. (This would block the application of (2*) of the schema of the Overdetermination Argument to everything a human causes.) The Overdetermination Argument would also cease to threaten us if we exercised 'downward causal control' over our atoms. (This would block (1*)'s application to humans.)

The next chapter argues, independently of the Overdetermination Argument, that humans cause things that their parts do not. It also argues that, as a result, humans have

non-redundant downward causal control over their constituent atoms. So if the arguments of the next chapter succeed, we are safe from the Overdetermination Argument. And in Chapter 5, I'll argue that we are safe from the considerations of Chapter 2.

But note that even if (contrary to fact!) the arguments of the next chapter failed, one could still deny that the Overdetermination Argument eliminates human organisms. For one could *tollens* instead of *ponens,* concluding that, because we exist, we either cause things not caused by our parts or exercise causal control over our parts.

One might accept these claims about a human's causal powers as the price of one's own admission into Being. And one might do so while consistently endorsing the Overdetermination Argument's application to baseballs and statues. For a metaphysics that attributes these sorts of causal powers to human persons is more plausible than one that attributes such powers to baseballs or statues. (For more on this point, see Chapter 6's (§III) discussion of free will.) But it would be nice to have independent confirmation of the truth of this 'top-down' metaphysics of humans. And we shall have it.

4 Surviving Eliminativism

Throughout the previous chapters the assumption that we are human organisms has been in the background. This assumption is now about to take centre to make a few remarks in support of it. I am (all too) aware that my remarks won't persuade everyone. They aren't intended to. They are meant to show only why the assumption that we are human organisms is so natural.

I think that when you look in a mirror, or down at your hands, you can actually see yourself. And when you hold your child, you do exactly that—hold the child himself or herself—and not some stand neither seen nor held. These claims also imply that human persons are neither mental states nor akin to software. Such software—as opposed to the hardware on which it 'runs'—could be neither seen nor held; similarly for mental states.[44]

Some think that persons are physical objects co-located with organisms. Such a view allows persons to be seen and held.

[44] Objection: we see a non-physical person by seeing her body. Response: this sense of seeing a person is secondary, parasitic on the primary sense in which we see her body; but I say we see persons in the primary sense.

But, as we have seen, there are problems with co-location generally. The co-location of a person and a human organism would, in fact, be especially problematic. After all, because the organism would comprise all the same atoms in all the same arrangements as the person with whom it would be co-located, it looks as if they would have all the same thoughts.[45] And so this sort of co-location seems to imply that none of us can tell whether he or she is the organism who thinks falsely that it's a person or, instead, the person who thinks this truly. (See Chapter 2, §IV; see also Carter 1988; van Inwagen 1990, n. 45; and Olson 1997.)

We are not co-located with human organisms. But we can be seen and held just as those organisms can. So I conclude that we just are, are identical with, human organisms. This does not imply the clearly false claim that *being a person* is the same property as *being a human organism*. Nor does it imply that we have both (or either) of those properties essentially. Nor am I claiming that human persons have the persistence conditions traditionally attributed to organisms.

Let me elaborate on this last point. It is not because I endorse a 'biological criterion of personal identity' that I identify each person with an organism. Rather, I make that identification because I oppose co-location. There is no co-location. So there is exactly *one* thing where we truly believe there to be a human person and a human organism (and a human body). Obviously, this implies that the person is identical with the organism (is identical with the body). But it does not—at least not obviously—imply that we have biological persistence conditions. For example, our currently being organisms might be consistent with our later undergoing the

[45] One could resist this inference by insisting that the organism has no mental properties even though the person does. This would imply that an object's mental properties did not supervene on its atoms' features and interrelations. Such non-supervenience is one possibility suggested, but not entailed, by the arguments of §II below (and suggested by different arguments in Rea 1997*a* and Shoemaker 1999*a*, *b*).

gradual replacement of our cells by circuits until we are no longer biological entities.[46]

One might object that, because I reject folk ontology and thereby deny certain deliverances of common sense, I have no right to assert confidently that you can see yourself in a mirror or that you can hold your child. After all, the objection concludes, these assertions are themselves nothing more than bits of common sense.

I defend some surprising ontological conclusions. But this does not imply that I am entitled to *no premisses whatsoever.* And surely I ought to proceed with premisses that seem true rather than with premisses that do not. The bits of common sense noted above seem true. So I proceed with them as premisses.

My opposition to folk ontology does not—all by itself—preclude my reasonably endorsing a claim simply because that claim seems right. So there is nothing to the objection just considered. But there is a potentially more substantial objection in the neighbourhood. This is the objection that the intuitively attractive claims that suggest that we are organisms are refuted by the arguments that refute folk ontology. Or more to the point—since we might as well face the most direct threat to my ontology—let us consider the objection that belief in our existing and being human organisms is thus refuted. So we need to ask whether the arguments of earlier chapters, which give us good reasons to deny the existence of statues and baseballs, give us good reasons to deny our existence as organisms.

This chapter and the next answer that question. Chapter 5 responds to threats to our existence (and our being organisms) inspired by Chapter 2. And this chapter explains how we avoid

[46] There is a sense in which humans have persistence conditions: we would survive some adventures (a walk in the park), but not others (a walk on the sun). But there are no criteria—no informative necessary and sufficient conditions—for our identity over time (Merricks 1998*b*; see also my 2000*b*). So, even though we are organisms, there are no *biological* criteria for our identity over time.

being eliminated by Chapter 3's Overdetermination Argument, even though we are composite organisms.

Now one might object that, even setting aside the Overdetermination Argument, Chapter 3 poses a problem for the claim that we exist. For recall that one's prima facie justification for believing that baseballs exist is undermined once one realizes that all of one's sensory experience regarding alleged baseballs would be the same whether caused by statues or merely by atoms arranged baseballwise (Chapter 3, §III). And so one might object that our justification for believing that we human organisms exist is likewise undermined.

In response, what ordinarily justifies our believing that we exist is not completely overdetermined by—and so is not undermined by—atoms arranged humanwise. One's evidence for one's own existence, and so for the existence of at least one human organism, is not straightforwardly sensory or even obviously causal. So even if our atoms did overdetermine all that we caused, our atoms might not wholly account for—and so might not undermine—some of our reasons for believing in ourselves.

Moreover, as I shall argue in this chapter, our atoms do not overdetermine all that we cause. We humans are not causally redundant. So our atoms may not even account for the sensory experiences we cause, at least not all by themselves and without relying somehow on us. It could be, for example, that we non-redundantly cause our atoms to cause those experiences.

So the epistemic points made in Chapter 3 do not threaten the claim that we exist. But what of the Overdetermination Argument? Note that that argument eliminates baseballs because, if they existed, they would be at best causally redundant. The Overdetermination Argument cannot eliminate us because—as this chapter will argue—we are not causally redundant. In brief, the argument of this chapter proceeds as follows. Our having conscious mental properties does not supervene on what our parts are like. We cause certain effects

by having such properties. Because our causing those effects is appropriately independent of what our parts are like, our parts do not cause those same effects. So we are not wholly causally redundant. So we survive the Overdetermination Argument.

I. Step One

The following is the Step One argument or, for short, 'Step One':

(1) There is some intrinsic property F such that:
  (*a*) An object's existing and being F is not, of metaphysical necessity, implied by the existence and intrinsic properties of, and spatiotemporal and causal interrelations among, that object's constituent atoms.
  and
  (*b*) Humans cause things in virtue of (existing and) being F.

(2) If (1) is true, then there is some property F such that a human's causing effect E in virtue of (existing and) being F does not *all by itself* give one a reason to believe that that human's constituent atoms cause E in virtue of their intrinsic properties and spatiotemporal and causal interrelations.

Therefore,(3)

There is some property F such that a human's causing effect E in virtue of (existing and) being F does not *all by itself* give one a reason to believe that that human's constituent atoms cause E in virtue of their intrinsic properties and spatiotemporal and causal interrelations.

The Step One argument is clearly valid. The only question is whether its premisses are true. In this section, I'll defend

premiss (2). I shall then defend premiss (1). Step One is but a first step towards saving us from the Overdetermination Argument. The latter part of this chapter completes the task.

Suppose that an object's constituent atoms (in virtue of having the relevant features) did fix, of metaphysical necessity, that object's causally relevant properties. Then it would be tempting to think that that object's causal powers could not possibly outstrip the powers of its atoms. It would be tempting to think that that object could cause only what its atoms also cause. Now I have already opposed the idea that what a composite causes can be *analysed as* what its parts cause (Chapter 3, §III). Nevertheless, one could still insist that, if, necessarily, an object's causal powers do not 'float free' of the powers of its parts, then there is some *a priori* or *conceptual* reason to think that the object causes only what its parts (such as its atoms) cause.[47]

But suppose, instead, that some causally relevant features of a composite object did 'float free'—in the sense specified by premiss (1) of Step One—of what its atoms were like. In such a case, the object's causal contribution would be in that sense independent of what its atoms were like and so, presumably, independent of what its atoms cause. In such a case, there seems to be no reason to think that what that object causes is also caused by its atoms.

If an object's causal powers are not fixed of metaphysical necessity by its parts, then we have no conceptual or a priori reason to conclude that what the object causes is caused by its parts. But, one might worry, suppose we have other reasons. For example, suppose we knew that it was a law of nature that

[47] What a composite causes cannot be analysed in terms of what its parts cause. So perhaps even if the powers of an object do not 'float free' of what its parts do, there is no conceptual reason to think that the object causes only what its parts cause. The British Emergentists (e.g. Mill 1843; Alexander 1920; Broad 1925; Morgan 1923), accepting that emergent properties supervened on—did not 'float free' of—microproperties, might have endorsed such a line. I think the line I defend in the text, focusing on properties that do 'float free', is more compelling.

an object composed of atoms causes only what its atoms cause. Or suppose we knew, in a particular case, that an object was caused by its atoms to cause an effect.

These suppositions are irrelevant to the point I'm trying to make here, in this section. For that point is only that (2) is true. That point is only that if a composite object's (existing and) being F is not fixed of necessity by its atoms, then that object's causing E by way of being F does not *all by itself*—unaided by, for example, the knowledge that the object's atoms caused it to cause E—provide any reason to believe its atoms cause E.

At least, it provides no such reason if being F is an *intrinsic* property. The existence of an object with certain *relational* properties is not, necessarily, implied by the intrinsic features and spatiotemporal and causal interrelations of its constituent atoms. Yet suppose that an organism's existing and having a particular relational property is so fixed by the relevant features of both the atoms that compose it and also the atoms that fill its environment.[48] In such a case, one might argue that all those atoms (in virtue of having the relevant features) cause whatever the organism causes in virtue of having that relational property. Suppose such an argument were successful. Then the atoms that compose that organism would, in virtue of the relevant features, 'partially cause' (alongside the atoms in the organism's environment) the effect in question. And to partially cause is to cause.

At any rate, suppose that an object's constituent atoms do *not* fix of metaphysical necessity what is both *intrinsic* and *causally relevant* about the whole. Then the whole's causing an effect gives us no reason—all by itself—to think those atoms cause that same effect. Thus premiss (2) of the Step One argument is true.

[48] I assume that if an *intrinsic* property is not fixed, of necessity, by its atoms, it is not so fixed by its atoms plus the atoms that fill its environment. See n. [49] below.

[49] Objection: Although conscious mental properties are intrinsic and C false, the existence of a conscious object globally and microphysically supervenes. Replies: (1) External factors might *cause* an object to have intrinsic features. (The potter causes a vessel to have a certain shape.) But intrinsic properties, just by their very nature, *supervene* either locally or not at all. (2) I have argued (1998*a,* §4) that if the exemplification of an intrinsic property fails to supervene locally, then it fails to supervene globally. (3) Even setting (1) and (2) aside, this objection is totally unmotivated. It will become obvious why I say this once we see, as I shall argue below, that even if no conscious mental properties are intrinsic, differences in consciousness fail to supervene on intuitively relevant microphysical differences. We shall see that, no matter what one says about the intrinsic nature of either being conscious or its supervenience base, no intuitively attractive thesis of global microphysical supervenience is available.

The above reasoning in support of (2) relied on the idea of an 'intrinsic property'. So let me conclude this section with some comments on this idea. Intrinsic properties are, by and large, those properties that an object can exemplify even if that object and its parts (if any) are the only objects that exist. This is a mere 'mark' of *being intrinsic*. It is not an analysis. If we tried to turn this mark into an analysis, it would have to be complicated to avoid some counter-examples. For example, *being the only object in the universe* bears the mark, as does *being a sphere that is not a proper part of another sphere*. But excluding funny cases like these, this mark seems to get things right and so it is useful. *Being oblong* comes out as intrinsic, since it is possible that the only object in the universe be oblong. But *being three feet from a dog* is not intrinsic; it is impossible that the only existing object be three feet from a dog.[49]

It is easy to find a loose-fitting mark of *being intrinsic*. An airtight analysis is something else altogether. Indeed, it is safe to say that there is no non-controversial analysis on the market, save the platitudinous one that to be intrinsic is to be nonrelational.[50] I don't want the Step One argument held hostage to the fortunes of any controversial analysis. So I'll rely, instead, on what is least controversial and most trustworthy: the platitude, particular clear examples, and the mark of *being intrinsic*.

I shall not endorse any (non-platitudinous) analysis of *being intrinsic*. But I must reject one. Consider the claim that *being intrinsic* is analysed as being a property whose exemplification by an object supervenes on the intrinsic features of, and (spatiotemporal

[49] This last example might need a bit of tweaking, since some suppose that, if space were very tightly curved, a lonesome dog could be three feet from itself! Nevertheless, the 'mark' captures something important in our intuitive understanding of *being intrinsic*. After all, purported analyses of *being intrinsic* often take the 'mark' as their starting-point. (See Kim 1982; Vallentyne 1997; Langton and Lewis 1998).

[50] And even the platitude needs to be qualified since relations to oneself or one's parts can be intrinsic properties (see Francescotti 1999).

and causal) interrelations among, the things that compose it. This analysis, if correct, would undermine the Step One argument. For, assuming that *supervening on* entails *being, of metaphysical necessity, implied by,* it would render premiss (1) false on the grounds that, trivially, no intrinsic property could fulfil condition (*a*).

But premiss (1) is, at least so far, safe. For this alleged analysis, employing a claim about the *intrinsic* features of parts in its analysans, is circular. And so we ought to reject it *as an analysis*. It remains an open question, however, whether it is true that an object's existing and having intrinsic features always supervenes on that object's parts in the way the failed analysis claims. I defend a negative answer to this question below.

One might try to cure the failed analysis of its circularity by deleting the right-hand-side occurrence of 'intrinsic'. However, the analysis thus amended is vulnerable to obvious counter-examples. For example, my being three feet from a dog supervenes on a certain atom of mine being three feet from a dog. Thus the amended analysis implies that *being three feet from a dog* is intrinsic.

II. Conscious Mental Properties and Premiss (1*a*)

Consider the many conscious and subjective (no wide content!) mental properties characteristic of a healthy, awake, adult human person. These properties bear the 'mark' of being intrinsic. After all, most theists believe that God might never have created; they believe there is a possible world that contains only God. Their belief implies that there is a possible world that contains just a single entity with many conscious mental properties. This implication is not rendered incoherent by the nature of such properties. (If it were, presumably, someone would have argued for atheism along these lines.)

If you have no taste for theology, consider the solipsistic hypothesis that I—an entity with many conscious mental properties—am all that exists. While surely false, this hypothesis is not rendered incoherent simply by the nature of those properties. (Contrast this with the hypothesis that I am the only existing object and am three feet from a dog.)

Many conscious mental properties bear the 'mark'. And an object's having such properties is not a matter of its being related to (or failing to be related to) some other thing.[51] Thus I conclude that many conscious mental properties are intrinsic. And so we can show that there is a property that satisfies clause (*a*) of premiss (1)—we can show that 'premiss (1*a*) is true'—if we can show the following: the existence of an object with an intrinsic conscious mental property is not entailed by the intrinsic properties and spatiotemporal and causal interrelations of that object's constituent atoms. Indeed, I shall show something a bit stronger. I shall show that the existence of an object with enough intrinsic conscious mental properties to have a rich mental life (for short: 'a conscious object') is not thus entailed. Showing this will show that (1*a*) is true.

So (1a) is true if the following is false:

> *Consciousness (C).* Necessarily, if some atoms $A_1 \ldots A_n$ compose a conscious object, then any atoms intrinsically like $A_1 \ldots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \ldots A_n$, compose a conscious object.

Suppose P is a normal and conscious human being. Suppose that P accidentally slices off her left index finger (i.e. her atoms arranged left-index-fingerwise) and thereby 'shrinks'. Suppose

[51] One might object that each conscious mental property is 'maximal', i.e. is analysed (in part) as being had by something that is not a proper part of an object that has that same property. Maximal properties, like *being a sphere which is not a proper part of another sphere,* bear the mark of being, but fail to be, intrinsic. This objection is best replied to below (§III), once its motivation is made clear.

that at the very first instant at which P has lost her left index finger, the atoms that then compose her remain just as they were (intrinsically and in all their spatiotemporal and causal interrelations) immediately before amputation. This implies—assuming C for *reductio*—that, just as those atoms compose a conscious object (P) after amputation, so they composed a conscious object before amputation. Name that latter object 'the finger-complement'. The pre-amputation fingercomplement is not identical with P. (Proof: P had a part, a left index finger, that the finger-complement lacked.) So before amputation, if C is true, there were *two* conscious entities, P and the finger-complement, sitting in P's chair and wearing P's shirt. But there was exactly one such entity. So C is false (cf. Merricks 1998*a*). So premiss (1*a*) of Step One is true.

My *reductio* of C relies on the claim that, before amputation, there were not two conscious entities (P and the fingercomplement) sitting in P's chair, wearing P's shirt. The claim that there really were two conscious entities—indeed, because they have equally rich mental lives, two persons—wearing P's shirt and sitting in P's chair leads to even greater absurdities. For if there was such an object as the conscious fingercomplement, it seems there was also a conscious toothcomplement, thumb-complement, toe-complement, and so on. And as it goes for P and her complement of complements, so, presumably, it goes for all of us. But this is false. Indeed, it's simply incredible. (It might even lead to our elimination; see Unger 1980.) There is not a mighty host of conscious, reflective, pain- and pleasure-feeling objects now sitting in my chair, now wearing my shirt, now thinking about metaphysics.

Some will say that 'there are many conscious beings now wearing my shirt' is ordinarily false, even given the mighty host. For they hold that, when conscious beings overlap sufficiently, we 'count them as one' (see Lewis 1976, 1993). This does not directly address the point at hand. For my claim is not that the truth of the *sentence* 'there are many conscious

beings now wearing my shirt' is absurd. It is rather that it is absurd that there are many (non-identical) conscious beings now wearing my shirt and thinking my thoughts.

My argument against C claimed that P survives the loss of a finger. And of course I think people can survive losing a finger. But, for the record, this particular argument can accommodate even the mereological essentialist. For this argument requires only that *some conscious being or other* (not necessarily P) exists after finger amputation. And surely *someone* is there. This, conjoined with C, implies that there was a conscious pre-amputation finger-complement. But there was not. For the existence of a conscious pre-amputation finger-complement leads to an unacceptable multiplication of persons.

III. Objections to the Defence of Premiss (1*a*)

The above argument assumed that, when P's finger is removed, the rest of her atoms remain unchanged in their intrinsic features and interrelations. But, one might object, this assumption is clearly false. Remove the finger, and, for example, blood starts clotting.

The argument against C need not involve anything so large as a finger. Imagine instead that *one atom* in P's finger is instantaneously annihilated. It seems plausible that, at the first instant that the atom fails to exist (or—if there is no 'first instant'—at some instant very shortly thereafter), the atoms that then compose P have not yet reacted to the change. And we can then show that C implies, absurdly, that both P and the atom-complement exist and are conscious.

Nor does it matter if, as a matter of contingent causal fact, the remaining post-annihilation atoms would react instantaneously to one of their kin's annihilation. The argument against C requires only that the following two claims are possible

(and compossible): after one of P's atoms is annihilated, a person exists composed of all the atoms, save the annihilated one, that originally composed P; at the very first instant that the annihilated atom ceases to exist, the atoms then composing a person are (in the relevant ways) just as they were at the preceding moment. Given C, these two claims lead to absurdity. So C is not necessarily true. So C, which purports to be a necessary truth, is false.

The argument against C—so another objection begins—requires that post-amputation P and pre-amputation finger-complement are exactly alike at the atomic level. But given perdurance (or four-dimensionalism), P's finger amputation implies only that the temporal part P has right after amputation is atomically just like the temporal part the finger-complement had right before amputation. That is a far cry from P and the finger-complement being exactly alike in the intrinsic features and interrelations of all their constituent atoms. So if persons are four-dimensional, the above argument against C fails.

In response, even if four-dimensionalism were true, we should still reject C. More carefully, we should still reject the *four-dimensionalist's version* of C. We need a four-dimensionalist version of C here because four-dimensional persons might be composed of only the proper temporal parts of atoms rather than (entire) atoms. For if at one time (as we would normally say) atom A was a part of P and at another A existed but was not a part of P, then the perdurantist says that the whole of A was never a part of P. Instead, only a proper temporal part of A was ever among P's parts.

Recall that we targeted C for refutation in order to support premiss (1*a*) of the Step One argument. So, in figuring out how to recast our target in terms congenial to perdurance, the four-dimensionalist's gloss on (1*a*) should be our guide:

(4D1a) There is some intrinsic property F such that an object's existing and being F is not, of metaphysical

necessity, implied by the existence and intrinsic properties of, and spatiotemporal and causal interrelations among, that object's *constituent atomic temporal parts*.

We can show that (4D1*a*) is true—and that our arguments therefore have purchase on the four-dimensionalist—just so long as we can show that the following is false:

4D Consciousness (4DC). Necessarily, if some atomic temporal parts $A_1 \ldots A_n$ compose a conscious object, then any atomic temporal parts intrinsically like $A_1 \ldots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \ldots A_n$, compose a conscious object.

Suppose that P is a four-dimensional person who lives exactly eighty years and is then instantaneously annihilated. Suppose further that (in the same world) another person, P*, is for the first eighty years of her life microphysically intrinsically just like P, although she outlives P by a decade. In other words, the atomic temporal parts that P* has for the first eighty years of her life are exactly like (in intrinsic features and causal and spatiotemporal interrelations) the atomic temporal parts that wholly compose P. 4DC implies that the atomic temporal parts that P* has for the first eighty years of her life compose a conscious object just like P. But they do not compose a conscious object at all, lest there be at least two conscious things where we know there is exactly one. So 4DC is false.

My argument against 4DC assumed that there are not at least two conscious beings where we know there is exactly one. Many perdurantists will protest. One reason that they will do so is that they tend to believe that a person has a 'temporary' property at a time—a property that can be gained or lost—only if that person's temporal part at that time has that property (see Lewis 1986*a*: 202 ff.). Conscious mental properties are

temporary. Thus perdurantists will hold that there are at least two conscious beings—the person and her stage—whenever and wherever we know there is exactly one. That's not an objection to my argument against 4DC; it's a *reductio* of perdurance.

(There is more we could say here. For example, I could add that four-dimensionalism renders you unable to know whether you have existed for more than a day, since you cannot tell whether you are a person or, instead, its today temporal part. Objection: 'I' refers to the 'most inclusive' entity, thus to the person; so the thought 'I have been alive more than a day' is true, whether thought by person or stage.[52] Reply: Whether or not that thought is true, if things seem to the stage just as they seem to the person, you can't tell whether you are a day-long stage or a lifelong person (see Chapter 2, §IV).)

We considered two objections to my C-denying defence of premiss (1*a*). The third and final objection, unlike the first two, concedes the falsity of C. It is that knowing that C is false would not justify endorsing premiss (1*a*) of the Step One argument. For, according to this objection, the conscious mental properties at play in the above argument are not intrinsic. Therefore premiss (1*a*) does not follow from the denial of C. Ironically, the best reason to deny that conscious mental properties are intrinsic seems to be the very point I've been labouring to defend: the falsity of C. Let me explain.

C is false. So if the relevant mental properties are intrinsic, then whether atoms compose a conscious object does not supervene on microphysical doings (i.e. on intrinsic properties of, and causal and spatiotemporal interrelations among,

[52] This objection—if we interpret it so as to imply that there is exactly one person where we know there is exactly one—implies that there is an entity that is the *single* 'most inclusive'; but—for a variety of reasons, including how they typically deal with vague identity over time—most perdurantists deny that any entity is such that it is determinately the 'most inclusive'.

microphysical entities).[10] But suppose those mental properties were extrinsic, relational. Then whether atoms compose a conscious object might supervene on the microphysical doings in *and around* the object. At least, the falsity of C does not preclude this sort of supervenience. So we must choose between two claims. The first is that many conscious mental properties are intrinsic (enough of them so that having only intrinsic conscious mental properties could constitute having a rich mental life). The second is that whether atoms compose a conscious person supervenes on the microphysical. One might object that the second claim is more compelling than the first.

My response to this objection begins by focusing on the intuition that motivates it. That intuition, I presume, is that atoms' composing a conscious object supervenes not only on microphysical doings, but on doings that are intuitively *relevant*. Our imagined objector would not be pleased to learn, for instance, that whether certain atoms compose a conscious object supervenes on how atoms light years away from that object are arranged. By the same token, she should be dismayed that whether there is a conscious being composed of certain atoms supervenes on whether that being is next to an atom—not in a brain but—in *a left index finger*.

Reflecting on the contrast between atoms arranged P-wise—recall that P is a conscious person—and atoms arranged P's-atom-complementwise shows something important. It

shows that we hope in vain when we hope that differences in whether atoms compose a conscious object always supervene on intuitively relevant and significant microphysical differences. Even if all conscious mental properties were extrinsic, whether certain atoms compose something conscious supervenes—if at all—on paltry and seemingly irrelevant microphysical detail.[53]

Some will object that the microphysical detail is not irrelevant nor its paltriness problematic. Their objection is rooted in the idea that being conscious is 'maximal'. We find an idea like this in Michael Burke's (1994, n. 21) claim that 'our concept of . . . a thinker' is maximal. And Harold Noonan (1999) notes that this sort of idea threatens reasoning of the sort I defend above. Here I shall focus on the version of this idea, and the resultant objection, defended by Theodore Sider (1999*b*).

To begin to understand Sider's view, consider his notion of *pseudo-consciousness*. According to Sider, there is one difference between being pseudo-conscious and being conscious. Being conscious precludes, but being pseudo-conscious does not preclude, being part of a 'larger' conscious entity—thus the 'maximality' of being conscious. So to be conscious is to be pseudo-conscious *and* to fail to be part of a larger conscious entity.[54] Crucially, there is *no phenomenological difference* between the conscious and the merely pseudo-conscious. A similar point applies at the level of individual conscious mental properties. The merely pseudo-conscious only 'pseudo-feel'

[53] One might object that, if conscious mental properties are intrinsic, we face a 'mysterious correlation'. Consider how differences in forming a conscious being are *correlated with* the relational differences between the atoms arranged P-wise and those arranged finger- and atom-complementwise (see Hawley 1998). I concede the correlation is mysterious. But this is no threat to the intrinsic nature of conscious mental properties. For even if they were extrinsic, the correlation would remain mysterious (see Merricks 1998*c*).

[54] That is, P's being conscious implies that P is not part of larger entity that has the numerically same thoughts as P. Sider's account is not meant to preclude, for example, billions of conscious entities' composing—like so many cells—some conscious being.

pain—but this just means they have all the phenomenology of feeling pain and are part of a bigger pain-feeling entity.

On Sider's account, conscious mental properties are *not intrinsic*. (They misleadingly bear the 'mark' because they are maximal.) So there is no need to explain how his account undermines my defence of (1*a*) of Step One. It's obvious. But I should explain how Sider hopes to make trivial microphysical details—such as whether someone has a particular atom as a part of her finger—intuitively relevant to whether one is conscious. On his view, what it *means* for something to be conscious rather than (merely) pseudo-conscious is that it is *not* part of a larger—larger even by a single atom—conscious object. Whether something is conscious (as opposed to merely pseudo-conscious) can then, potentially, amount to something extremely trivial. (It need not amount to anything as momentous as whether it has a rich phenomenology.) And so, given Sider's view, it is quite plausible that whether something is conscious could supervene on a trivial difference of the sort noted above.

Note that we can have most of Sider's benefits without his 'maximal' view of consciousness. We could simply assert that both P and the atom-complement exist and are conscious. We could then conclude that there is *no* difference in phenomenology, consciousness, or existence between P and the atom-complement supervening on their trivial atomic difference.

This 'multiple conscious entities' response gives us most of the benefits of Sider's position because, in a way, it *is* Sider's position. Sider's position is merely a notational variant of this response. Yet his notational variation allows him to say that the sentence 'there is exactly one conscious entity wearing my shirt' is true. Thus Sider might claim that he does not run afoul of our intuition that there is only one conscious entity in my shirt, just so long, that is, as this is 'properly interpreted' as a semantic intuition about the truth of sentences containing the word 'conscious'. (Sider will presumably say that other

intuitions about 'consciousness'—such as that many conscious mental properties are intrinsic and that whether one has them cannot turn on whether one has a certain atom in one's finger—are really intuitions about pseudo-consciousness.)

But I think the fact that Sider's position is merely a notational variant of the 'multiple conscious entities' response, dressed up with some new semantics for 'conscious', is reason enough to reject it. For the problem with the multiple conscious entities approach never had anything to do with semantics. The problem has always been that response's ontology. (Recall similar complaints in §II about 'counting many as one'.) It is, I say, both false and incredible that there are many beings now wearing my shirt who have all the *phenomenology of consciousness*. Fiddling with the words—calling all but one of these beings 'merely pseudo-conscious'—makes this neither more true nor more believable.

Moreover, Sider's semantics for 'conscious' seems mistaken because it has the implausible result that one cannot tell whether 'is conscious' describes oneself. In the material mode, Sider's semantics implies that one cannot tell whether one is conscious. For, given Sider's view, there seems to be no way to tell whether one is the 'maximal' object—and so genuinely conscious—or just one of the countless wannabes. (I guess they only pseudo-wanna.)

At any rate, I shall proceed on the assumption that the multiple conscious entities view, and notational variants thereof, is false. As a result, no matter what we say about the intrinsicness of conscious mental properties, we must reject the intuitively attractive picture of supervenience. We must reject the picture according to which differences in whether atoms compose a conscious object supervene on relevant and correspondingly significant microphysical differences. Once the *intuitive* claim about supervenience is gone, there is little to be said for the remnant: every difference in whether atoms compose a conscious person supervenes on a microphysical difference, but

that microphysical difference can be paltry and intuitively irrelevant.

As noted above, at most one of the following theses is true: first, many conscious mental properties are intrinsic; second, whether atoms compose a conscious object supervenes on microphysical doings. No intuitively compelling version of the second thesis is available. Add to this the 'mark' of being intrinsic, the possibility of a lonesome conscious entity, and the fact that whether something is conscious seems not to be a matter of its being (or failing to be) related to some other entity. All of this supports the first thesis. Many conscious mental properties—enough such that the having of them would constitute having a rich mental life—are intrinsic.

IV. Step One Again

Recall the first premiss of the Step One argument:

(1) There is some intrinsic property F such that:

(*a*) An object's existing and being F is not, of metaphysical necessity, implied by the existence and intrinsic properties of, and spatiotemporal and causal interrelations among, that object's constituent atoms.

and

(*b*) Humans cause things in virtue of (existing and) being F.

If (1*a*)'s being made true by some conscious intrinsic mental properties is to lead to a defence of premiss (1) as a whole, then humans must cause things in virtue of having some of those properties. In Chapter 6 (§I) I shall address the threat of mental epiphenomenalism. But for now I assume that we cause things in virtue of having intrinsic conscious mental properties. So I conclude that premiss (1) is true. I have already

defended premiss (2). And Step One is clearly valid. Step One is complete.

Before taking Step Two, I want to explain why Step One cannot be transformed into a first step towards saving, for example, statues from elimination. The key here is that few properties can undergird both (1*a*) and (1*b*) (or an appropriate analogue of (1*b*)).

Consider quidditative, non-qualitative, properties, such as, for example, *being statue S* (or *being Trenton Merricks*). One might plausibly argue that an object's quidditative properties are both intrinsic and fail to be entailed by the intrinsic features and spatiotemporal and causal interrelations of its constituent atoms. Thus such properties might satisfy the demands of (1*a*). But they do not seem to satisfy the demands of (1*b*). For merely quidditative properties don't seem to do any causal work. (Hence some of the suspicion about their very existence.)

One might object that *being statue S* allows a statue to cause the *true* belief 'here is statue S', which a statue otherwise like S could not cause. Thus one might conclude quidditative properties do causal work. This is the wrong conclusion. Statues exactly alike—except that one is S and the other not—in exactly the same situations would cause exactly the same beliefs. If one of those beliefs were 'here is statue S', statue S could *make that belief true,* another statue not. But 'truthmaking' is not causal. This is implied by the bare possibility of a causally inefficacious property P, since it could play a role in making 'P is exemplified' true.

Let us turn to qualitative properties. Few of these can satisfy the demands of Step One. Take, for example, a statue's mass. This is causally relevant and so could satisfy the appropriate analogue of (1*b*). But the statue's 'finger-complement'—suppose S is a statue of Moses pointing upward—will differ from the statue in mass. So we cannot show that the statue's mass satisfies (1*a*) simply by mimicking § II's argument against C.

Consider now the property of *having a mass of at least M*. Suppose the statue and its finger-complement (exist and) share this feature. Suppose further that that feature is genuinely causally efficacious. Even given all this, it is hard to see how we could get a compelling argument going in support of the appropriate version of premiss (1). For unlike the claim that there are two conscious beings wearing my shirt, the claim that there are two objects upon the pedestal, each of which has a mass of at least M, is not absurd.

There is one final challenge we need to consider. Suppose composition is, in general, 'primitive'. That is, suppose that whether atoms *compose something* is not necessarily fixed by their intrinsic features or causal and spatiotemporal interrelations. In other words, suppose that whether *some object exists* that is composed of those atoms is not fixed by their intrinsic features, etc. This supposition combined with Step-One-style reasoning implies the following about, for example, a 100 pound statue: the statue's causing the scale to register '100' gives us no reason to think that its constituent atoms cause the scale to register '100'. So, one might conclude, just so long as Step-One-style reasoning is sound to begin with, we are on our way to saving 100 pound statues (and every other alleged composite) from the Overdetermination Argument.

This challenge could get off the ground only if we had reason to believe that, in general, composition was primitive. We do not have reason to believe this. We do, on the other hand, have reason to believe that C is false. And one explanation of why it is false is that, in cases involving humans at least, composition is primitive. But having one such potential explanation of C's falsity is a long way from having reason to believe composition in general—and so in the case of a statue—is primitive.

In what follows, I shall argue that Step One opens up a way for *us* to avoid being eliminated by the arguments of Chapter 3. But analogues to it do not seem to offer a similar escape for

(alleged) non-living macrophysical objects. For some of our properties, conscious mental properties, satisfy the demands of *both* clauses of premiss (1) of the Step One argument. I can think of no properties that statues (or baseballs or rocks or stars) would have that plausibly manage to satisfy both clauses of the appropriate analogue.

V. Step Two

The Step One argument is sound and so the following, its conclusion, is true:

(3) There is some property F such that a human's causing effect E in virtue of (existing and) being F does not *all by itself* give one a reason to believe that that human's constituent atoms cause E in virtue of their intrinsic properties and spatiotemporal and causal interrelations.

Establishing (3) is only a first step towards saving ourselves from the Overdetermination Argument. A second step is necessary because (3) does not, on its own, show that we should deny or even withhold belief on the claim that everything a human organism causes is also caused by its constituent atoms. Nor does it suggest that a human is causally relevant to what its constituent atoms cause. Thus (3) alone does not show that either (1*) or (2*) of the schema of the Overdetermination Argument (Chapter 3, §IV) cannot be applied to all that we supposedly cause. And so it does not, on its own, save us from that argument.

(3) fails to block the application of (1*) or (2*) to us. This is because, in part, (3) is restricted to only some of the properties of the relevant atoms. (Obviously, atoms have properties that are not intrinsic and stand in relations that are neither causal nor spatiotemporal.) So one question we need to ask is

whether a human's causing something in virtue of being conscious gives us a reason, all by itself, to think her atoms cause that same thing in virtue of some of their features *other than those singled out by the Step One argument.*

There is, so far as I can tell, only one sensible approach to answering this question affirmatively. One could claim that a human's causing something in virtue of having a conscious mental property automatically gives us a reason to think her atoms caused that same effect *in virtue of composing a conscious person who causes that effect.* But relying on such a claim to support the Overdetermination Argument as applied to humans renders humans immune to elimination. For it is (trivially, obviously) impossible for the atoms in question to stand in the (allegedly) causally relevant relation without composing a person and so without a person existing.

Moreover, I deny that atoms cause anything in virtue of being interrelated by *composing a conscious person who causes something.* In the closing seconds of the sixth game of the NBA Finals, Michael Jordan stole the ball from Karl Malone. This caused the Chicago Bulls to win the 1998 championship. As much as I'd like a share of the credit, I don't think the victory was overdetermined as a result of my being such that Jordan steals the ball from Malone. Nor, for similar reasons, do I think that atoms cause E by being such that the person they compose causes E.

But set that claim aside. Still, we should conclude the following:

(4) Suppose a human causes effect E in virtue of (existing and) having a conscious mental property. This does not, all by itself, give one a reason to believe that that human's constituent atoms cause E in virtue of any of their properties or interrelations other than those which would, trivially, entail the existence of a conscious person (who causes E in virtue of having a conscious mental property).

(4) alone will not save us from the Overdetermination Argument. This is because (4) has inherited (3)'s other significant qualification. (4), like (3), leaves open the possibility that there is some *other* reason, other than the person's causing the effect, to think her atoms cause that effect (and do so in a way that does not trivially imply that the person exists and causes something). To complete Step Two, and to complete our rescue, we need to address this qualification.

The possibility that (3) and (4) leave open is an actuality. To begin to see why I say this and also why—perhaps surprisingly—this is no problem for my position, consider the following. A person shatters a 'window' by throwing a 'baseball'. Such a shattering is caused by the person and also by the atoms arranged baseballwise. But this shattering is not thereby overdetermined. This is because the person causes the atoms to strike the window, thus making her causally relevant to whether they shatter the window. And, as should be clear from Chapter 3, overdetermining causes must be causally *irrelevant* to one another.

This example illustrates that, given that a person causes E in virtue of having a mental property, we may very well have reason to think that some atoms cause E. It is obvious why this example so plausibly provides a case in which some atoms cause an effect also caused by a person: the person is not causally irrelevant to the atomic causes; the atomic causes and the person's deciding are part of the same 'causal chain'. The real oddity would be if some atoms caused an effect, an effect that was caused by a person's having a mental property, from *outside the causal chain* that included the person's having that mental property.

For these reasons, and in light of (4) above, I conclude that, barring the odd case where we have good evidence to believe overdetermination has occurred, the following is true:

(5) Suppose a human person causes an effect E in virtue of (existing and) having a conscious mental property. Then

> there is, with one possible exception, no reason to think that that same effect is caused by *any* atoms in such a way that the person is causally irrelevant to those atoms causing that effect. The possible exception is that there may be reason to think those atoms cause E by way of composing a person who causes E in virtue of having a conscious mental property.

Sometimes my *deciding* to do such and such is what causes the atoms of my arm to move as they do. Presumably my so deciding won't ever be the *only* cause of their moving. There will also be a cause in terms of microphysics or microbiology, in terms of nerve impulses and the like. But at some point in tracing back the causal origin of my arm's moving (if it is intended), we will reach a cause that is *not* microphysical, that just is the agent's *deciding* to do something.

Someone might object to (5) because she believes that *every* physical effect has microphysical causes to which nonmicrophysical causes—such as mental causes—are causally irrelevant. (We shall return to a thesis like this in Chapter 6 (§I), under the name of 'Microphysical Closure'.) She might add that I too ought to endorse this objection to (5). After all, she might remind us, I insist that everything a *baseball* would cause—if it would cause anything at all—would be caused by its atoms. Moreover, I insist that the baseball is causally irrelevant to what those atoms cause. She might think that if I insist on such claims regarding baseballs, I ought to accept them regarding humans as well.

The response to this objection is to remind us of the overall strategy of this chapter. This chapter argues that the existence of some objects with causally relevant properties (namely, objects with conscious mental properties) does not supervene on microphysical doings. Because of that, I have argued, we should say that some of what those objects cause, in virtue of having those properties, lack microphysical causes.

Yet I endorse the exceptionless existence of microphysical causes with respect to the effects of (alleged) baseballs. This is, in part, because we have no compelling argument for the claim that, if baseballs existed, their existing and having some causally relevant property would fail to supervene on the microphysical. That is, no analogue of premiss (1) of Step One for baseballs (or statues) is as compelling as the original premiss (1) itself (§IV).

Moreover, recall the arguments in Chapter 3 (§II) for the claim that the baseball's atoms shatter the window. One such argument was that every atom arranged baseballwise causes something, and when what one of them causes is added to what each of the others causes, the 'sum' is the shattering of the window. And a similar point holds for everything the baseball seems to cause. But it does not seem that, for example, when what one of my atoms does is added to what each of the others does, the 'sum' is my consciously deciding.

But I am happy to agree that (5) has the virtue of being falsifiable. For (5) could—at least in principle—be undermined by empirical evidence for the claim that every physical effect has a microphysical cause to which non-microphysical entities are causally irrelevant. Such vulnerability to empirical refutation is something I share with my opponents here. For the philosopher who *accepts* the claim that every physical effect has microphysical causes (to which non-microphysical causes are causally irrelevant) takes no less of an empirical gamble than do I.

Understatement for effect: Microphysicists have not yet causally explained every physical event. How one bets on whether they will ultimately do so should turn on, among other things, the arguments. Because of the argument of this chapter up to this point, I conclude that (5) and its implications for the completability of microphysics are true.[55]

[55] My position on conscious mental states might remind some of the British Emergentists. But they seem to explain *being emergent* in *epistemic* terms. They seem to say that a property is emergent just in case we cannot *predict* its appearance, given perfect knowledge of its supervenience base (e.g. Broad 1925: 52 ff.). I have nowhere explained *being a non-redundant causal property* in epistemic terms. Moreover, paradigmatic emergentists were committed to mental states' supervening on the physical; I am not.

(5) is true. As I have emphasized before, unless there is some reason to think overdetermination has occurred, we should assume that it has not. In light of this and the truth of (5) we should, except for the odd cases where we have evidence that some funny coincidence has occurred, endorse the following:

(6) Suppose a human person causes an effect E in virtue of (existing and) having a conscious mental property. Then E is caused by her constituent atoms *only if* that person is causally relevant to whether those atoms cause E (or those atoms cause E by way of composing a person who causes E in virtue of having a mental property).

We can now show that the schema of the Overdetermination Argument cannot be applied to the effects a human causes by having conscious mental properties. And so it cannot be applied to *every* effect caused by a human person. Thus we have the resources to show that the Overdetermination Argument does not render us humans completely epiphenomenal. This, in turn, prevents us from being eliminated as a result of the reasoning that, in Chapter 3, eliminated baseballs and other inanimate macroscopica. (In what follows, I shall assume that whatever we have said about atoms can also be correctly said about a person's parts at other levels of decomposition.) Here's the schema:

(1*) Object O—if O exists—is causally irrelevant to whether its parts $P_1 \ldots P_n$, acting in concert, cause effect E.
(2*) $P_1 \ldots P_n$ cause E.
(3*) E is not overdetermined.

Therefore,(4*)

If O exists, O does not cause E.

If a human person, in virtue of having a conscious mental property, causes an effect E, either her parts $P_1 \ldots P_n$ cause that effect or they do not. If $P_1 \ldots P_n$ do not cause that effect, then (2*), applied to this case, is false. If this is how things go, then obviously the schema of the argument cannot be applied to everything we allegedly cause. The argument cannot render us completely epiphenomenal.

Suppose, on the other hand, that our imagined human's parts $P_1 \ldots P_n$ *do* cause the effect she caused by having a conscious mental property. Then given (6) above, one of two results follows. The first result is that her having the mental property is causally relevant to $P_1 \ldots P_n$ causing E. This result indicates that (1*), applied to this case, is false. And this result implies that the Overdetermination Argument fails to render us completely epiphenomenal.

There is a second result consistent with (6). (6) allows that $P_1 \ldots P_n$ could cause E simply by composing a person *who causes E*. But this cannot be a first step towards arguing that the person is epiphenomenal. For anyone who endorses (2*) applied to our imagined person on these grounds must—because he is thereby committed to the person's causing E—reject (3*) applied to the present case. And so the argument again fails to render us causally inefficacious.

The argument fails, but—so I say—not because (3*) applied to the case at hand is false. For I deny that there is, simply as a matter of logic, widespread overdetermination. So, for example, I oppose the claim that for any object O that causes an effect, every other object causes (and thus overdetermines) that effect simply by being such that O causes it. And likewise—to repeat a point stated earlier in this section—I deny that a person's atoms (or other parts) cause an effect simply in virtue of composing a person who causes that effect.

Although we survive the Overdetermination Argument, we do not do so on the grounds that we overdetermine some of our effects, effects caused by our atoms being such that we cause those same effects. If we did survive by way of the falsity of (3*) applied to some of what we cause, then one might be tempted to revisit the question of whether the baseball might not, after all, overdetermine the shattering of the window. But that's not how we survive; so there's no temptation to revisit.

Even setting this last point aside, we can see that the Overdetermination Argument cannot be adapted to show that every effect a conscious human person is alleged to cause is not actually caused by her. Thus Chapter 3's reasons for eliminating statues and other inanimate macroscopica do not give us a reason to eliminate humans. Step Two is complete.

VI. On What Composite Objects Exist

Composite objects that cause things that their parts do not redundantly cause can resist the eliminative sweep of the Overdetermination Argument. We humans—in virtue of causing things by having conscious mental properties—are causally non-redundant. So the Overdetermination Argument fails to show that we do not exist. So I conclude that we do. For we should assume that we exist unless we are shown otherwise. Any conscious composita presumably survive the Overdetermination Argument just as we do. So I conclude that dogs and dolphins, among other animals, exist.

Perhaps some sort of 'biological anti-reductionism' is true. Perhaps all organisms—not just the conscious ones—cause things not redundantly caused by their parts. If so, all organisms can resist elimination via the Overdetermination Argument. But I don't know whether 'biological antireductionism' is true. And so, for example, I don't know

whether there are trees or ants. Similarly, I don't know whether the atoms of physics—if they existed—would cause some effects not overdetermined by their parts. If not, then I conclude there are no such things.

If there are no atoms, I have to recast many of my earlier claims. I'd have to talk about things other than atoms 'arranged statuewise', 'working in concert to shatter windows', and so on. The nature of these 'other things' does not matter. They could be simple. They could be composite. They could be quarks, leptons, gluons, and photons. They could be molecules. (Suppose non-redundant causal properties 'emerge' when things are arranged moleculewise.) Or they could be things for which we have no name.

I have been careful not to build into eliminativism the claim that simples exist (see Chapter 1, §I). My ontology is, of course, consistent with simples' being the ultimate constituents of the material world. It is also consistent with matter's infinite divisibility, with *every* material thing having proper parts. In light of the Overdetermination Argument, I tentatively conclude the following. If matter is infinitely divisible, then there are new levels of causal powers descending *ad infinitum*. That is, there are new levels of objects with causal powers that don't merely reproduce the powers of those objects' parts.

So suppose that all composite material objects have non-redundant causal powers. Trivially, if there are simples, they have non-redundant causal powers (and so there is nothing mysterious about non-redundant causal powers per se). Then, for material objects, *to be is to have non-redundant causal powers*. Earlier I assumed that all macrophysical objects have causal powers (Chapter 3, §IV). By contrast, what the arguments of this chapter and the previous suggest is that every material object not only has causal powers, but has *non-redundant* causal powers.

Material objects must be causally non-redundant. Or so I conclude for now. I would have to abandon this conclusion if

there were a good argument for the claim that matter is infinitely divisible, and that, moreover, at many or all microscopic 'levels' the wholes cause only what their parts cause. This claim would imply a lot of systematic overdetermination at the microscopic level. And I concede that it would ameliorate such overdetermination by macroscopica. But, strictly speaking, systematic microscopic overdetermination is consistent with my arguments for eliminating statues and baseballs. For even if systematic overdetermination were forced upon us in one place, we might still want to resist it in others, accepting it only where we have no choice. And, as we have seen, we do have a choice when it comes to the overdetermination allegedly wrought by statues and baseballs and other inanimate macroscopica.

VII. Conclusion

Human organisms do not dodge the Overdetermination Argument on a mere technicality of which baseballs, for example, cannot avail themselves for some intuitively irrelevant reason. Rather, human organisms have non-redundant causal powers and so can exercise downward causation. Baseballs, on the other hand, would not—even if they existed—have non-redundant causal powers or exercise downward causal control over their parts. This deep, fundamental difference between the powers of human organisms and the powers of alleged baseballs (and statues and rocks and stars and so on) makes all the difference with respect to the Overdetermination Argument.

This chapter raised a couple of issues that it did not adequately address. The first is mental epiphenomenalism. The second revolves around those things a human causes, but does *not* seem to cause directly by having a conscious mental state.

(Imagine, for example, that I am thrown through a window. My atoms seem to shatter the window; I seem to shatter the window; and, as a result, the shattering of the window seems overdetermined.) These issues will be addressed in Chapter 6, where I continue the exploration of mental causation and the causal powers of human organisms begun in this chapter.

But discussions of our causal influence would be, at best, merely hypothetical if it turned out that we did not really exist. And so, before embarking on such discussions, it is important to show that we are not eliminated. In this chapter I showed that we are not eliminated by the arguments of Chapter 3. In the next I'll show how we survive the arguments of Chapter 2.

The arguments of this chapter strengthen those of the former. The arguments of Chapter 3 might have seemed *too* powerful, eliminating all actual and possible composite physical objects, holding up a standard for existence too lofty for any composite object to meet. But now we see that this is not so. The arguments of the preceding chapter are, it turns out, discriminatory. They give us a reason to deny the existence of some supposed composites but not others. And thus the plausibility of those arguments is increased.

5 Considerations in Favour of Eliminating Us?

Chapter 2 raised some considerations in favour of eliminating inanimate macrophysical objects such as statues. The obvious question—a question which arose again and again during that chapter—is whether those same considerations support eliminating human organisms. This chapter will answer that question, addressing each section of Chapter 2 in turn. I shall also explain how these considerations bear on microscopica, some of which compose us and others of which are arranged statuewise.

I have already argued that statues *et al.* do not exist (Chapter 3). And I've defended the existence of human organisms (Chapter 4). Nevertheless, I shall not invoke the non-existence of statues in contradistinction to our existence in responding to objections below. For such responses would presuppose my ontology. But I intend my responses—when combined with the considerations of Chapter 2—to constitute a nonquestion-begging defence of that ontology.

I. Persons and the Water in the Pool

I claimed that eliminativism about 'the water in the pool' (where 'the water' is taken to be a single large material object) was neither particularly radical nor obviously at odds with folk ontology. I then suggested that alleged statues seem relevantly similar to 'the water in the pool'. So I concluded that the plausibility of eliminating the water lends at least some plausibility to eliminating statues (Chapter 2, §I).

The obvious objection is that these considerations lend just as much plausibility to eliminating persons as they do to eliminating statues. In response, there are at least three reasons why eliminating persons is *less plausible* than eliminating statues, given that eliminating the water is not obviously inconsistent with folk ontology. To begin to understand the first of these three, recall that the reason that folk ordinarily believe in statues—sensory experience—turns out not to provide a good reason for believing in statues (Chapter 3, §III). Recall further that no similar problems afflict the reasons one ordinarily believes in oneself, owing to, among other things, the first-personal (non-causal-perceptual) nature of some of those reasons (Chapter 4, introduction). Knowledge of one's own existence is thus comparatively more secure than supposed knowledge of the existence of a statue.

Further support for this claim can be found by waxing Cartesian. Supposed knowledge of macroscopic objects is more likely to be mistaken (e.g. because one is dreaming) than is supposed knowledge of one's own existence. Moreover, since I cannot know anything without existing, to know that a statue exists, I must exist. So there is something fishy about supposing that I know a statue exists but do not know that I exist. There is nothing likewise piscatory in the claim that I know that I exist but not that a statue does.

Now one might object that we are, after all, waxing *Cartesian*.

The considerations invoked here, this objection continues, might secure knowledge of one's own existence, but they do not allow one to conclude that one is an organism. I agree. These considerations tell us nothing about whether (or—*contra* Descartes—not) we are organisms. But I am exploring whether the assumption that we exist and are organisms can hold up under the sort of scrutiny that forced us to abandon the assumption that there are statues and baseballs. So I can rightly start off assuming that if we exist, we are organisms (see Chapter 4, introduction). And so reasons to think one exists are—given that assumption—reasons to think a human organism exists.

Supposed knowledge of our own existence is privileged in comparison to that of the existence of statues. Therefore, the claim that belief in one's own existence is mistaken is less plausible than the claim that belief in the existence of statues is mistaken. Thus we have our first reason that—given our assumption that each human person is a human organism—eliminating at least one human organism is less plausible than eliminating statues, even given that eliminating the water is not obviously at odds with folk ontology.

This first reason supports the claim that human organisms in general—as opposed to only oneself—exist. After all, it would be odd if one's own atoms arranged human-organismwise compose an organism, but all other atoms thus arranged fail to do so. Indeed, 'arranged human-organismwise' roughly means of some things that they are such that, if humans exist, they compose one (see Chapter 1, §I). So if *any* humans—including oneself—exist, atoms arranged human-organismwise *always* compose a human.[56]

[56] But this is not as straightforward as it might seem. Atoms arranged statuewise are not arranged material-objectwise. For if they were, and if there were material objects—and of course there are—then they'd compose one (see Ch. 1, §I); but they don't compose anything; so they aren't arranged material-objectwise. With this in mind, note that if one were the only human organism, no atoms other than one's own would be arranged human-organismwise—even if atoms other than one's own seemed (because of their spatiotemporal and causal interrelations) to be thus arranged.

Suppose the H_2O molecules in the pool do not compose an object. Nevertheless, they are somehow 'united'. This unity, which *ex hypothesi* is not that of composing something, results in part from how we think and speak, from our grouping the molecules together in thought and referring to them collectively with a single expression, 'the water in the pool'. (Other factors play a role in 'uniting' the molecules, such as, for example, their being more or less clumped together.) If all this is right, it is at least somewhat plausible that atoms arranged statuewise are united not by composing something but, instead and in part, by how we speak and think.[57]

A parallel claim about atoms arranged humanwise is not equally plausible. It is not equally plausible that atoms arranged humanwise do not compose anything—and so do not compose a human—but are united, instead and in part, by how we think and speak. For this implies, given that we are humans, that the 'unity' among atoms arranged humanwise is a result of how we think and speak, even though we do not exist. This is incoherent.

Some might object that, although of course nothing can result from how we think and speak if we never exist, this is beside the point. For they will object that it is not we who 'unite' atoms by thinking and speaking; thinkings and speakings themselves do this. And, the objection continues, thinkings and speakings can exist without thinkers or speakers.

The claim that atoms arranged humanwise fail to compose a human but are united, in part, by thinkings and speakings implies something that parallel claims about statues and 'the water' do not imply. It implies the exotic (and perhaps

[57] A similar point was made in Ch. 1 (§I), when I said that the folk concept of *statue* plays a role in fixing whether atoms are arranged statuewise. But I don't want to rely on that discussion here. For that discussion awarded that role to the folk concept only on the assumption of eliminativism's truth.

demonstrably impossible[58]) claim that there are thinkings and speakings but neither thinkers nor speakers. So there is an expense to extending the eliminativist picture of 'the water' to include humans, an expense not incurred by extending that picture to include statues. This is the second reason that eliminating human persons is less plausible than eliminating statues, even given that eliminating the water is not obviously at odds with folk ontology.

Suppose a baseball causes the shattering of a window. Even so, we can account for the shattering without invoking the baseball. This is because some other things (some atoms) actually cause the shattering. Indeed, as should be familiar from Chapter 3, the baseball is at best causally redundant. But, as we saw in Chapter 4, persons are not causally redundant.

Because a person is causally non-redundant, a person causes effects that cannot be automatically accounted for in his or her absence. Because a baseball is causally redundant, every effect a baseball causes can be automatically accounted for in its absence. Thus eliminating persons has a cost that eliminating baseballs (and statues) does not. That cost is the need to fill the 'causal gaps' that persons—but neither baseballs nor statues—would leave behind if eliminated. This cost is the third reason that eliminating persons is less plausible than eliminating statues, even given that eliminating the water is not obviously at odds with folk ontology.

This section opened with the obvious objection that the plausibility of eliminating 'the water' makes eliminating us human organisms just as plausible as it makes eliminating statues. The three reasons given above—the special epistemic access each has to his or her own existence; the elimination of persons, but not statues, implying thinkings without thinkers; and the fact that, unlike persons, everything baseballs and

[58] For arguments against the Humean claim that thinkings and speakings do not require thinkers or speakers, see Strawson (1959, ch. 3), Shoemaker (1997: 139), and Lowe (1989; 1996: 25 ff.).

statues are alleged to do is automatically accounted for without them—show that the 'obvious objection' is mistaken.

We could rely on our three responses to the 'obvious objection' in replying to *any* objection purporting that an argument supports eliminating statues only to the extent that it supports eliminating human persons. For given these three responses, any argument will have a harder time making our elimination seem plausible than making the elimination of statues seem plausible. But other than to note this fact here, I shall not make use of it when responding to other alleged considerations in favour of our elimination.

Perhaps there are no simples. Perhaps, as a result, there are composite microscopica. And so, having answered the 'obvious objection' regarding persons, I shall consider it as applied to the composite atoms of physics. But before doing so, I want to remind us that the point here is not to defend atoms per se. It is, rather, to show that I am not committed to simples by showing how I handle composite microscopica. As always, atoms are merely our stand-in for whatever the composite microscopica might turn out to be (Chapter 1, §I; Chapter 4, §VI).

It is not clear that considerations regarding the water support eliminating atoms to exactly the same extent that they support eliminating statues. Suppose, for example, that atoms cause things not caused by their parts. Then a relevant difference between atoms and statues is that they, but not statues, cause things not automatically accounted for on the supposition of their non-existence.

Or suppose, instead, that there was a good reason to believe in composite microscopica that was not a good reason to believe in statues. For example, suppose there were a sound argument for the impossibility of simples. That argument would rule out eliminating all composite microscopica. But it would not rule out eliminating statues. And thus that argument itself would draw a distinction, regarding the plausibility

of elimination, between alleged statues and some composite microscopic objects or other.

Maybe there is no such argument. And maybe everything each alleged composite microscopic object would cause (if it existed) would be caused by its parts. If so—and only if so—then Chapter 2's reflections on the water render eliminating composite microscopica as plausible as eliminating statues. That's fine with me.

II. Persons and the Sorites Game

We assumed for *reductio* that Michelangelo's *David* exists. We then imagined annihilating one of *David*'s atoms after another, asking God, after each annihilation, whether *David* still exists. It seemed to follow that, on the occasion of some single crucial annihilation, God's answer would change from 'yes' to 'no'. One could deny that this follows, but only by paying one or another hefty price. One price is metaphysical vagueness and *David*'s vague identity. The other price is the early-in-the-Game co-location of myriad kindmates, the *David*-candidates.

In the service of metaphysical economy, I concluded that if *David* existed, God's answer would change from 'yes' to 'no' upon the annihilation of a single atom. But it is implausible that God's answer would so change. So, I claimed, we should say that *David* never existed in the first place (Chapter 2, §II).

I responded to a number of objections to the argument just outlined. But I left one unanswered. That objection begins by noting that—given my allegiance to metaphysical economy—I can accept neither that a human's identity can be metaphysically vague nor that each human person is co-located with many others. So I must conclude that a single atomic annihilation would make all the difference in whether a human exists, were we to exist in the first place to be subject to the Sorites Game.

Thus, this objection concludes, our Sorites sorties against *David* eliminate it if and only if they eliminate us human organisms.

I agree with the objection that, were we to play the Game with humans, a single atomic annihilation would make 'all the difference'. But this is not a reason to eliminate humans.[59] A similar result regarding statues is, however, a reason to eliminate them. Defending this difference between humans and statues—thus responding to the objection just noted—is the task of the rest of this section.

We have already seen that a human's existing and being conscious is not necessarily implied by any features of that human's constituent atoms (other than those—like their standing in the *composing a conscious person* interrelation—that do so trivially) (Chapter 4). One of two results follows. First, a human's existing is not entailed by the relevant features of his or her atoms (see Chapter 4, §IV). Or, second, although a human's existing is so entailed, his or her being conscious is not. In the following two paragraphs I'll suppose, as may be the case, that the former is true.

A human's existing is not entailed by his or her constituent atoms' having the relevant features. Therefore, those atoms' *composing a human* is not entailed by their having the relevant features. This implies that there could be two microindiscernible histories—each concluding with the same atoms having the same features—yet only one of which concludes with its respective atoms composing a person. This in turn implies that there could be two identical micro-histories differing in personal identity. And so—here's the punch line—

[59] I think identity is 'what matters in survival'. But Parfit denies that what matters in survival can turn on a microphysical triviality, such as, presumably, the annihilation of a single atom (1984: 239). My response to Parfit would begin by charging him with conflating what identity 'consists in' with what it 'turns on'. (See Merricks 1999*c*.)

there can be differences in personal identity *not grounded* in microphysical differences.[60]

If such differences can be microphysically ungrounded, it is no surprise that they can be correlated with trivial microphysical differences, differences like the annihilation of a single atom during the course of the Sorites Game. And so the Sorites Game, played with humans, results in nothing implausible. (Wheeler 1986: 345 ff. defends a similar point.) This line of argument cannot, however, be adapted to render plausible the results of the Sorites Game when played with *David*. For folk ontology, with its bottom-up metaphysics for statues, forbids differences in statue identity over time that fail to be grounded in microphysical differences.

Thus we have one reason that the Game may not count against our existence to the extent that it counts against *David*'s. But I don't want to rely too much on this reason. For, as explicitly noted above, it requires us to make a supposition that I have not fully defended here. Happily, there is another reason—the primary reason—to draw the necessary distinction between *David* and humans that does not require this supposition.

In presenting this reason, I begin with the fabled 'last straw' that broke the camel's back. It is surely plausible that a *single straw* can make all the difference in whether the camel's back breaks. This is plausible *whatever* one says about vagueness. For this has *nothing to do with vagueness*.

(I assume (somewhat artificially) that the camel's back's breaking is not itself vague. I assume that the camel stands tall until the last straw snaps its back, dropping it to the ground.

[60] Cases of fission give a more decisive reason to conclude that differences in personal identity need not be grounded in microphysical differences (see Merricks 1997*a*, esp. 172–4). One might object that considerations in the text—and those in Merricks (1997*a*) —imply that personal identity fails to supervene locally on the microphysical, not that it fails to supervene globally. But the failure of local supervenience here entails a failure of global supervenience: see Merricks (1998*a*, §4); Ch. 4, n. 10; and especially Merricks (forthcoming *a*, §VI).

Those uncomfortable with this pretence should run the argument to follow with a trap door which remains shut until, upon the addition of the final straw, it drops open.)

I hope it is obvious that some cases in which a trivial difference of one sort generates a non-trivial difference of another sort have nothing to do with vagueness. I also hope it is obvious (given our pretence) that the case of the camel and the last straw is such a case. What is less obvious is what, exactly, makes a single straw's breaking a back thus different from a single hair's balding a man.

I tender the following suggestion. When the final straw is added, something occurs that is altogether different from the mere addition of that final straw. Perhaps the suggestive but hard-to-define idiom of constitution might help. If the loss of a single hair renders a man bald, his becoming bald is somehow *constituted by* his losing that last hair. There is—in some sense—nothing more to his becoming bald than his coming to have the resultant number (and distribution, etc.) of hairs. In contrast, when the addition of a single straw breaks a camel's back, the camel's having a broken back is in no way *constituted by* its having that straw and its predecessors placed upon it. It is false that the camel's having a broken back is—in any sense—nothing more than its bearing that number of straws.

Now consider what a Cartesian dualist might say, were we to play the Sorites Game with a human organism. She might say that at some point during the Game the soul would flee the body. She might say its fleeing would be correlated with (and presumably caused by) the annihilation of that last crucial atom. Yet our Cartesian dualist is not thereby endorsing a controversial interpretation of a case of vagueness. The body's ceasing to be ensouled here—even if not empirically verifiable—is, with respect to vagueness, not at all like a person's going bald with the removal of the last crucial hair.[61] For

[61] Dean Zimmerman made this point at the UNC-Greensboro Symposium on Ontology, Greensboro, NC, Apr. 2000.

there is no sense in which a soul's leaving a body is constituted by, or is nothing more than, the annihilation of the key atom.

We human organisms cause things our parts do not cause. Thus when one of us goes out of existence as a result of the annihilation of a single atom, that result is altogether different from—is *not* constituted by—that atomic annihilation. For causal forces disappear from the world that can't be treated as, in any sense, 'nothing more than' the disappearance of the causal forces of the annihilated atom.

So here is what I say. Because human persons cause things their parts do not cause, a human person's ceasing to exist's being correlated with a single atomic annihilation has as little to do with vagueness as would—if Cartesian dualism were true—the soul's leaving the body's being so correlated. And both of these 'correlations' have as little to do with vagueness as does a camel's back's breaking being correlated with the addition of a single straw.

If *David* existed, it would cause nothing not caused by its parts working in concert. Thus the removal of the atom that would remove *David* would not make a difference that is appropriately 'in addition to' the removal of the atom itself. Note that the only causal differences—the only differences even possible—would be the differences resulting from the loss of the work of the last atom itself. For the statue's causal powers, if it has any, are equivalent to those of the atoms that compose it. (That's why folk ontology implies systematic overdetermination.)

Of course, the atomic annihilation—itself an event—that results in *David* 's ceasing to exist (if *David* exists in the first place) is numerically distinct from the event of *David* 's ceasing to exist. One might think that this undermines the point of the preceding paragraph. But one would be wrong. A man's losing that last crucial hair is not identical with his becoming bald. (He could, after all, become bald without losing *that* hair.) Yet neither losing the final hair nor annihilating *David* 's watershed

atom brings about—in addition to the loss or annihilation—something 'altogether different', at least not in the sense of bringing about an additional change in the causal powers at work in world. In each case, the occurrence of the 'small' change seems somehow to constitute the occurrence of a 'large' one.

The claim that the addition of a single straw would 'make all the difference' with respect to a camel's back breaking does not imply anything controversial about the nature of vagueness. Nor is that claim itself implausible. It seems exactly right. The claim that a single atomic annihilation would 'make all the difference' with respect to a person's existence, like the claim about the last straw, implies nothing controversial with respect to vagueness.[62] More importantly, once its resemblance to a parallel claim about the last straw is noted, we can see its plausibility.

The claim that a single atomic annihilation 'makes all the difference' in whether *David* exists implies something controversial about the nature of vagueness. It is like the claim that the removal of a single hair 'makes all the difference' in whether one is bald. And so that claim about *David* is open to debate. (Indeed, it is worse than the epistemicist's claim about baldness, for even defenders of the episte-linguistic approach to vagueness should baulk at it (Chapter 2, §II).)

It is not obvious that the Sorites Game tells against composite microscopica. For it is not obvious that we *can* play the Sorites Game with entities like atoms. The Game can be played only with objects that have a large number of parts, parts that can be annihilated, seriatim, without thereby constituting or causing dramatic differences in the object's overall

[62] I am not defending an epistemic interpretation of the vagueness in personal identity during an episode of the Sorites Game. This is important because in Ch. 2 (§II) I argued that the best epistemic account—the 'episte-linguistic' account—requires multiple candidates of the sort that, if applied to cases involving personal identity, would lead to 'too many thinkers'.

state. (We would be unable to play the Game with *David* if, for example, it exploded when we annihilated a third atom.) Perhaps atoms do not have a large number of the right kind of parts. And so perhaps the Game cannot be played with atoms.

Or perhaps atoms—or whatever microscopic entities are down there—have non-redundant causal powers. Then it could be that the Sorites Game played with atoms would be like the (unpopular with camels) Straw Game, in which case the Sorites Game has no implausible implications for atoms. If, however, the Game could be played with atoms or other microscopic composita, and if those composita were causally redundant, then it would support eliminating them. But that's all for the good.

III. Statues, Lumps, and Persons

I argued that we should oppose co-location. I argued that we ought to deny that numerically distinct material objects could be composed of all the same parts at any level of decomposition (Chapter 2, §III, and the end of Chapter 3, §IV). I then argued that, if we deny that a statue is co-located with a lump of clay, we ought to conclude that neither statues nor lumps exist. For to say that, for example, the statue exists but not the lump is objectionably *arbitrary*. It is best, I argued, to deny the existence of lump and statue alike (Chapter 2, §III).

The worry at hand, of course, is that this reason to deny the existence of statues and lumps might turn out to be an equally good reason to deny *our* existence. For if we reject co-location, we cannot accept the co-location of a human organism with (for example) a lump of human tissue. And, one might worry, to say that human organisms exist, but alleged lumps of tissue do not, is objectionably arbitrary. Thus, one might conclude, the best course is to deny the existence of organisms *and*

lumps. Given that we are human organisms, this is to deny our existence.

To begin to understand my response to this objection, consider the epistemic parity of belief in a statue and belief in a lump of clay. The statue is allegedly composed of all and only the atoms which allegedly compose the lump. Thus the statue, if it existed, would cause all and only what the lump would cause, if it existed (Chapter 3, §IV). Add to this that the evidence one would have for either a statue's or a lump's existence would be ultimately causal–perceptual. It follows that evidence for the statue's existence is equally evidence for the lump's existence and vice versa. And because our evidence favours the existence of neither one over the other, it would be arbitrary for us to do so.

Each person has first-personal 'evidence' of his or her own existence. Such evidence is not evidence for the existence of an alleged lump of tissue. Thus such evidence makes it nonarbitrary to affirm one's own existence while denying the existence of one's alleged companion lump of tissue. And once one has concluded that one exists—and believes one is an organism—one has a reason to think other human organisms exist which is not also reason to believe in lumps of human tissue (or aggregates of humans' atoms or any other objects allegedly co-located with humans).

Moreover, we could run the Overdetermination Argument (of Chapter 3) applied to lumps of tissue (or aggregates, etc.). We could then note (relying on Chapter 4) that such an argument cannot eliminate us conscious human organisms. And so we'd have a reason to suppose lumps (and aggregates, etc.) do not exist that is not a reason to suppose human organisms do not exist.

I have suggested two reasons to favour human organisms over lumps of tissue. These amount to two responses to the charge that it is arbitrary to avoid co-location by denying the existence of lumps of tissue while affirming the existence of

human organisms. But we are not out of the woods yet. For both responses are vulnerable to a single objection.

Suppose we reject co-location of lump and organism. So we conclude that there is at *most* one thing composed of the atoms arranged lumpwise–organismwise. And suppose we agree that there is at *least* one thing so composed. (There are two reasons to agree to this. First, we have first-person knowledge of its existence. Second, being conscious, it can't be eliminated by the arguments of Chapter 3.) But, so the aforementioned objection goes, our two reasons for believing that there is some such thing—and thus our two responses to the charge of arbitrariness—do not indicate whether that thing is an organism or a lump. They indicate only that that thing is conscious and has first-personal knowledge of its own existence. To insist on its being an organism rather than a lump, the objector concludes, is arbitrary and unmotivated.

This is not the objection with which this section opened. That objection, modelled on Chapter 2, §III's consideration against the existence of statues, was that the least arbitrary response to the threat of co-location involves *denying* our existence. But the current objection does not intimate, in any way, that we do not exist. Rather, this objection is that—though we exist, are conscious, and have first-personal knowledge of our own existence—it is arbitrary to claim that we are human organisms rather than lumps of human tissue.

Even if this objection won the day, the revisions forced upon my ontology would not undermine its general 'spirit'. I could still defend my arguments against the existence of statues, baseballs, and other inanimate objects. I could still maintain that we are macrophysical objects. I could still argue that we survive the Overdetermination Argument by non-redundantly causing effects in virtue of our conscious mental properties. And so on.

But this objection will not win the day. For there is a good reason to think that we are organisms rather than lumps. A

lump of tissue, if there were such a thing, could not survive losing even a small bit of tissue. (Recall that lumps are defined in terms of their restrictive persistence conditions.) An organism, if there were such a thing, could survive losing a small bit of tissue. We can survive losing some atoms, and indeed some atoms arranged small-bit-of-tissuewise. So it is not arbitrary to say that we are organisms, not lumps (or aggregates).

In Chapter 4 (introduction) I distanced my claim that we are organisms from the claim that we have biological persistence conditions. And so one might charge that I have not defended our being organisms in the sense of 'organism' at issue in discussions of co-location—discussions in which sortal-relative persistence conditions loom large. One might then conclude that being an 'organism' in my sense does not preclude being a lump.

Part of this charge is correct. I do not avoid co-location by endorsing one set of controversial persistence conditions—such as 'biological conditions'—for human persons over another. And so my position is even less controversial than it might first appear to be. But part of this objection is mistaken. Although I do not call us 'organisms' in deference to biological persistence conditions, I do not apply that label capriciously. The label is appropriate for the following reasons, among others. We are physical things of human shape and size. We digest food and we breathe. We are members of the species *Homo sapiens*. And we can survive losing some atoms arranged small-lump-of-tissuewise; that is, we are not lumps of tissue.

Another objection focuses not on us human organisms, but on atoms. I reject co-location. So I deny that an atom is co-located with a numerically distinct lump of atom-stuff. But, one might object, it is arbitrary to say that the atom exists, but not the lump of atom-stuff. It is likewise arbitrary, the objection continues, to say the lump exists but not the atom. The least arbitrary position, our objector concludes, is to deny the existence of both atoms and lumps of atom-stuff.

I do not think anyone believes in both atoms and numerically distinct lumps of atom-stuff on the basis of intuitions about their incompatible persistence conditions (or, for that matter, on any other basis). I doubt there is anyone who is inclined to say, for example, 'Atoms can survive losing an electron *and* such atoms are wholly co-located with objects that cannot survive losing an electron'. There is no threat of co-location at the atomic level in the first place. And because the threat never arises, there's no need to respond to it, arbitrarily or otherwise.

But suppose, just for the sake of argument, that we found ourselves inclined to believe in atoms and in numerically distinct lumps of atom-stuff. Then I would agree that, *ceteris paribus,* the least arbitrary response would be to deny the existence of both sorts of entities. If, moreover—and contrary to fact—we faced a legitimate threat of co-location given *any* composite microscopica, I'd conclude that, *ceteris paribus,* the best response would be to deny the existence of all composite microscopic objects.[63]

But *cetera* might not be *paria*. For example, we might discover that simples are impossible. If so, then denying the existence of composite microscopica would not be the best response, all things considered, to the threat of co-location. Or suppose we had reasons to think that atoms cause things that their parts do not, but not so for lumps of atom-stuff. This would, given the arguments of Chapters 3 and 4, provide a non-arbitrary reason to avoid co-location by eliminating lumps of atom-stuff but not atoms.

[63] I make the standard assumption that simples face no threat of co-location. But consider a statue without proper parts. Could it survive squashing? Some might be inclined to say 'no'; the statue ceases to exist when squashed. Yet it seems that something survives: the simple lump which was co-located with the simple statue.

IV. Brains, Thinkers, and Persons

I argued that a human organism and his or her (alleged) brain cannot both exist, lest there be too many 'thinkers'. So we have two options. We could conclude that although there are atoms arranged brainwise, there are no brains. Or we could conclude that although there are atoms arranged human-organismwise, there are no human organisms. Either way, folk ontology is in trouble. Either way, we have a good example of things arranged F-wise but no Fs. Given the aims of Chapter 2, I stopped the argument there, merely noting that I believe in human organisms, not brains (Chapter 2, §IV).

Suppose someone chooses brains but no organisms rather than organisms but no brains. Suppose she thinks the 'thinker'—and so the person—is a brain, not a human organism. She thus objects to my ontology by denying that we are human organisms, indeed by denying that such organisms even exist.

This objection, like one in the preceding section, gestures more at friendly amendment than at wholesale refutation. For it grants that we exist and think. It grants that we are composite macrophysical objects. And it poses no threat to my arguments against statues nor to our surviving those arguments by, for example, non-redundantly causing things in virtue of our mental properties.

Nevertheless, I want to respond to this objection. Let us begin by considering the two strongest points in *favour* of the claim that each of us is a brain. First, it seems that one 'goes with the brain' in cases of brain transplant. The most direct way to underpin this widely shared intuition, one might add, is that one *just is* one's brain. Second, one might follow Descartes and claim that we are thinking things, but part ways with him by adding that the only thinking things are brains.[64]

[64] According to Parfit (1984, §98 and app. d), Thomas Nagel inclines towards the view that we are brains, and does so for this second reason (but see also Parfit's qualifications; 1984: 469–70). Elsewhere, Nagel (1986: 40) explicitly denies that he is a brain but claims his view on these matters could be expressed, 'with mild exaggeration', as the view that 'I am my brain'.

I agree that the view that we are brains is more plausible, *ceteris paribus,* than any view that denies either that one 'goes with the brain' or that each human person is a thinking thing. But I deny neither of these things. As already noted (Chapter 2, §IV), I think human organisms 'go with the brain' in 'brain transplant': we are 'whittled down' to brain size and 'grown' via the attachment of kidneys, lungs, etc. back to body size. And I agree that I am a thinking thing. Indeed, since on my view there is no brain (or soul or central nervous system), the only remotely plausible candidate for being the human thinker is the human organism.[65]

Although I say that we are human organisms and that brains do not exist, I can easily accommodate the two intuitions that best motivate the view that we are brains. Moreover, there are motivations for the 'organism view' that the 'brain view' cannot accommodate. These involve a return to commonsensical considerations akin to those noted at the start of Chapter 4, such as that I can see myself in a mirror. (I can't see my brain in the mirror.) Or that I can kiss my child. (I've never kissed a brain.) Or that I weigh more than 100 pounds. (That would be some brain.) And so on. Thus it is more plausible to say that we are organisms while eliminating brains than to say that we are brains while eliminating organisms.

[65] On the folk view, according to which brains exist, I think the brain is the best candidate for being the thinker. The organism seems to think only derivatively, in virtue of having a part that is the 'real' thinker. One could, of course, insist that human organisms are 'real' thinkers just like their brains are. But one is thereby vulnerable to the charge of 'too many thinkers'.

V. Conclusion

This chapter responded to challenges to my claim that we exist and are human organisms. These challenges were not chosen randomly. They mimicked Chapter 2's considerations in favour of eliminativism. Chapter 2 gives reasons to deny the existence of statues (and of some other alleged macrophysical objects) which do not, we saw in this chapter, lead to equally good reasons to deny our existence as human organisms.

Chapter 4's account of how persons avoid elimination by way of the Overdetermination Argument actually strengthened the Overdetermination Argument. For Chapter 4 showed that that argument's demands on composite objects were neither unrealistic nor impossible to satisfy. Persons satisfy them. Similarly, this chapter strengthens Chapter 2. For this chapter shows that Chapter 2 supports eliminating some material objects in a way that it does not support eliminating others.

Chapter 2 concluded on a modest note, emphasizing only the intelligibility of eliminativism and that eliminativism should be at least somewhat of a live option. We can now be less modest. Because of their discriminatory nature, Chapter 2's considerations offer considerable support for eliminativism.

This chapter and Chapter 2 constitute one strand of argument in support of my favoured ontology. A second, different strand of argument comprises Chapters 3 and 4. Thus we have two fairly independent and complementary defences of a statueless world populated by, among other things, us human organisms.

6 Mental Causation and Free Will

I RELIED on the causal kick of conscious mental states in defending our existence from the Overdetermination Argument (Chapter 4). Mental epiphenomenalism would undermine that defense. In this chapter I shall block what I see as the strongest argument for mental epiphenomenalism. I shall then respond to an overdetermination-based challenge to my metaphysics of human persons. Finally, I shall explore one way that metaphysics bears on free will.

I. The Exclusion Argument(s)

Jerry Fodor (1989: 77) says:

> if it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying. . . . if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world.

To put it mildly, the onus is on those who would deny mental causation. In other words, the onus is on the defenders of

mental epiphenomenalism. I shall focus on one line of argument for mental epiphenomenalism, the 'Exclusion Argument'. Versions of this argument have been discussed by, among many others, C. D. Broad (1925; cited in Yablo 1997), Norman Malcolm (1968), and, most recently and ably, Jaegwon Kim (1989*a, b*; 1993*a*; 1998: 37–8).

Here is my reconstruction of the Exclusion Argument. Assume for *reductio* that some physical event has mental cause M, that is, is caused by a person *in virtue of* her having mental property M. But *every* physical event (that is caused by an object's having a property) is caused by an object in virtue of its having a physical property. So the physical event that has mental cause M has physical cause P. M is causally irrelevant to whether P causes that event.[66] Thus any physical event caused by M is overdetermined. But there is no such overdetermination. Therefore, what we assumed for *reductio* is false.[67] And since this argument can be run for any alleged mental cause, we should conclude that *no* physical events have mental causes.[68]

I have incurred a special debt to address this argument. For in Chapter 3 (§III) I cited, in support of my own position, Kim's squeamishness about causal overdetermination, squeamishness he expresses while defending a version of the Exclusion Argument. Moreover, the Exclusion Argument closely resembles my Overdetermination Argument. So I need to show how I can consistently endorse the Overdetermination Argument while rejecting the Exclusion Argument's conclusion.

[66] Cause M is *causally irrelevant* to whether P causes an event just in case: M is not identical with P; M is not a 'partial cause' of the event alongside P; M does not cause P, nor does P cause M, to cause the event (see Ch. 3, §I).

[67] Note that the Exclusion Argument does not deny that mental *events* have physical effects. It denies that mental events have physical effects *in virtue of their mental properties*. Inspired by Brian McLaughlin (1989), we could say that the Exclusion Argument defends mental 'type epiphenomenalism', but not mental 'token epiphenomenalism'.

[68] Similar reasoning, based on the assumption that a mental state can be caused only by causing its subvenient physical base, rules out mental causes of mental states. C's falsity (Ch. 4, §II) should make us question this assumption.

The Exclusion Argument assumes that every physical event (that is caused by an object) is caused by an object in virtue of its physical properties, to which causing its having mental properties (if it has them) is causally irrelevant. It is this assumption—the principle of 'physical closure'—I shall oppose. As I interpret this principle, it implies that whatever a composite macroscopic object causes in virtue of its physical properties is also caused by atoms (or other microscopica) working in concert.

My interpretation of this principle is appropriate. For consider what macrophysical causes *without* corresponding microphysical causes would imply. A macrophysical object would have, in virtue of its physical properties, effects that are *not* caused by atoms working in concert. Such physical properties would thereby be 'emergent' and causally non-redundant (see Chapter 3, §II). And no self-respecting defender of the Exclusion Argument will allow that a macrophysical object has causally non-redundant physical properties.

She won't do so because allowing it would push the Exclusion Argument to the brink of begging the question. Allowing it puts her in the awkward position of accepting 'emergent' causally non-redundant properties as such, but not 'emergent' causally non-redundant mental properties, which would undermine her argument. The actual Exclusion Argument, as standardly defended, is not remotely question-begging. For it does *not* presuppose that while some properties have powers from beyond the realm of microphysics, mental properties do not. Rather, it insists that *no* properties have such powers. It insists that there are *no* non-microphysical properties that are causally non-redundant. It insists that 'completed physics' will ultimately provide 'bottom-up' causal explanations of everything physical.[69]

[69] Note that my argument to follow does not require that these anticipated explanations be in terms of things that are particularly small—for example, it could be that the explanations involve fields as opposed to quarks. What follows requires only that the promised explanations of the 'completed physics' will not be in terms of familiar macroscopic objects causing things in virtue of having familiar macrophysical properties.

Thus those who endorse the type of physical closure assumed by the Exclusion Argument should endorse the following:

> *Microphysical Closure*. Every physical event caused by a human in virtue of her having a mental property has microphysical causes to which non-microphysical causes—including that mental cause—are causally irrelevant.

Everyone who endorses (standard, 'non-emergentist') physical closure should (and probably does) endorse Microphysical Closure. So those who defend the Exclusion Argument ought to defend Microphysical Closure and, with it, the following 'Micro Exclusion Argument':

Assume for *reductio* that some physical event has mental cause M. But *every* physical event (that has a cause) has *microphysical causes* (or a microphysical cause). So every physical event that has M as a mental cause has microphysical causes. M is causally irrelevant to whether those microphysical causes cause that physical event. As a result, every physical event caused by M is overdetermined. But there is no such overdetermination. Therefore, what we assumed for *reductio* is false. No physical events have mental causes.

The Micro Exclusion Argument parallels the Overdetermination Argument even more closely than did the Exclusion Argument. It is, therefore, incumbent upon me to explain why I reject the Micro Exclusion Argument in spite of my devotion to the Overdetermination Argument. But before doing so, I want to note some important facts about the Micro Exclusion Argument.

The Micro Exclusion Argument is more of a threat to my ontology than is the Exclusion Argument. Defending mental causation is of interest here primarily because of mental

causation's role in preventing our elimination (Chapter 4). With this in mind, consider the Exclusion Argument's 'physical cause P', which supposedly overdetermines whatever M would cause. It seems plausible that P is supposed to be a *human's* exemplifying some non-mental property, such as having a brain in such and such a state. If so, then the Exclusion Argument implies that human persons exist. The Micro Exclusion Argument, on the other hand, undermines mental causation without providing any assurance at all of our existence. In this way, the Micro Exclusion Argument is more of a threat to my ontology than is its inspiration, the original Exclusion Argument.

Indeed, we should focus all our attention on the Micro Exclusion Argument, leaving the Exclusion Argument behind entirely. For the Micro Exclusion Argument charges that mental causation leads to unacceptable overdetermination *just because mental causation is not microphysical*. And so the Micro Exclusion Argument rules out mental causation if and only if that argument—or an obvious adaptation of it—rules out *all* non-microphysical causation. The Micro Exclusion Argument therefore rules out the sort of *physical* causation (macrophysical causation) the original Exclusion Argument uses to undermine mental causation. From all this it follows that if the Micro Exclusion Argument is sound, then the Exclusion Argument is unsound.

Those who endorse the Exclusion Argument should endorse the Micro Exclusion Argument. After all, the two arguments are alike in form and, as I have argued, physical closure (of the sort that motivates the Exclusion Argument) implies Microphysical Closure. But, as just noted, the Micro Exclusion Argument undermines the Exclusion Argument. So *no one* should endorse the Exclusion Argument.

I think this is a significant result. It is tempting to downplay this result, however, in light of the intimate connection between physical closure and Microphysical Closure. For one

might claim that because these doctrines (and their respective exclusion arguments) are so intimately related, we ought to take the Micro Exclusion Argument as merely clarifying, rather than supplanting, the original Exclusion Argument.

This would be a mistake. The Micro Exclusion Argument differs significantly from the Exclusion Argument. As noted, the Micro Exclusion Argument rules out non-microphysical causes of *all sorts,* not just mental ones. The Exclusion Argument never even hints at such a sweeping conclusion.

Moreover, the Exclusion Argument can be blocked in ways the Micro Exclusion Argument cannot. The most obvious response to the Exclusion Argument says whatever it must to identify the Exclusion Argument's mental property M with its physical property P (Kim 1989*b,* 1998). But this response fails even to suggest a reply to the Micro Exclusion Argument. For a human's mental property—such as her deciding to move her arm—is identical with neither the many properties exemplified by many microscopic entities (they're too many!) nor a single property exemplified by a single microscopic entity (it's too small!).[70]

Leave the Exclusion Argument behind. It is unstable. The Micro Exclusion Argument is another story. It is a stable, and serious, threat to mental causation and indeed to our very existence as organisms. It is at least as compelling as the standard Exclusion Argument. Moreover, it is invulnerable to the most obvious response (i.e. mental–physical property identity) to the Exclusion Argument. Thus the considerable interest

[70] Even panpsychists (such as, perhaps, Chalmers 1996: 293–301) should deny that the mental properties at play in the Micro Exclusion Argument—such as *deciding to move one's arm*—can be exemplified by a single atom.Note also that Yablo's (1992) response to the Exclusion Argument fails to engage the Micro Exclusion Argument. For neither many microphysical events (being too many) nor a single microphysical event (being too small) can be a *determinate* of a mental event. For similar reasons, Yablo's (1992) general line fails to engage the Overdetermination Argument, despite his remarks in nn. 3 and 5. A single baseball is neither a determinable nor a determinate of many atoms arranged baseballwise.

focused on the Exclusion Argument ought to be redirected towards the Micro Exclusion Argument.

The Micro Exclusion Argument is significant. But it is unsound. Recall that in Chapter 4 I argued that, except for the possibility of the occasional and odd causal coincidence, the following is true:

(5) Suppose a human person causes an effect E in virtue of (existing and) having a conscious mental property. Then there is, with one possible exception, no reason to think that that same effect is caused by *any* atoms [in virtue of any of their properties] in such a way that the person is causally irrelevant to those atoms causing that effect. The possible exception is that there may be reason to think those atoms cause E by way of composing a person who causes E in virtue of having a conscious mental property.

In the course of defending (5) I assumed that we cause things in virtue of being conscious (i.e. in virtue of having conscious mental properties). Thus one might object that relying on (5) and its defence, in responding to a challenge to mental causation, is question-begging. But this objection is mistaken. For the Micro Exclusion Argument is a *reductio*. It purports to show that if we assume mental causation, we end up with unacceptable overdetermination. My non-question-begging strategy to undermine that argument is to show that if we assume mental causation—as I did in my defence of (5)—we need *not* end up with unacceptable overdetermination.

(5) is true. Unless there is good reason to think otherwise, we should assume there is no causal overdetermination. So the moral of (5) is—or at least nearly is—the falsity of Microphysical Closure. (Recall my opposition to something like Microphysical Closure in Chapter 4, §V.) We might need the qualification of 'or nearly is' because (5) leaves open the following possibility. We might have good reason to believe

that an event caused by a person in virtue of her having a mental property is also caused by atoms in virtue of their composing the person *who causes that event by having a mental property*. Now I do not think that this is a way for atoms to cause things (see Chapter 4, §V, and §II below). But, for the sake of argument, let us suppose I am wrong.

Let us suppose that whenever one causes an event by having a mental property, that event is overdetermined by one's atoms in virtue of their composing a person who causes that event by having a mental property. Someone might insist that this saves Microphysical Closure. After all, they might note that Microphysical Closure places no restrictions on *how* microphysical entities act as causes, thus allowing them to do so by way of composing a person who causes something. This way of saving Microphysical Closure trivially entails that mental epiphenomenalism is false and thus that the Micro Exclusion Argument is unsound.

The move just suggested does not save the Micro Exclusion Argument. And, at best, it saves only the letter of Microphysical Closure. But I say that we should read Microphysical Closure in a way more true to its spirit. So I shall henceforth read Microphysical Closure as requiring 'pure' microphysical causes, that is, microphysical causes other than those that are obviously parasitic on non-microphysical causes. Thus read, Microphysical Closure is rendered false by (5). And because Microphysical Closure is false, the Micro Exclusion Argument is unsound.

In the first several chapters of this book I examined a domain in which I believe something relevantly like Microphysical Closure to hold, the domain of inanimate macrophysical objects such as statues and baseballs. I believe that, whenever something is allegedly caused by a statue or a baseball, then that something is also caused by atoms. And in the early chapters of this book, we discovered that the relevant variety of microphysical closure gets us nothing but microphysical objects.

That is, we discovered that domains in which every physical effect has ('pure') microphysical causes to which nonmicrophysical causes are causally irrelevant are domains with (barring overdetermination) only microphysical objects. We must resist systematic causal overdetermination. Thus in so far as one is committed to microphysical causes—as I am with respect to what is putatively caused by baseballs and statues—one is committed to eliminating macrophysical objects. Conversely, if one thinks there are *any* macrophysical objects of any sort, then one should reject the relevant variety of microphysical closure. So it should be no surprise that, given that humans exist, Microphysical Closure is false.

Note, for the record, that we could invert the argument of this section. Suppose (as I do) that we must deny Microphysical Closure to block the Micro Exclusion Argument and so to save mental causation. One could then conclude that because—lest it be the end of the world—mental causation occurs, Microphysical Closure must be false. And so humans, by having mental properties, cause things that are not overdetermined by their own atoms or by any other microscopica. We are now close to affirming the controversial sorts of claims defended in Chapter 4 regarding persons: persons cause things their parts don't cause and, as a result, exercise downward causal control over their parts. And so those committed steadfastly to mental causation can, without relying on the arguments of Chapter 4, save us from the Overdetermination Argument.

II. Causal Overdetermination Again

I have resisted systematic causal overdetermination at every turn. Yet one might charge that much of this book is misguided because systematic causal overdetermination is simply

unavoidable. Such a charge might be motivated by the idea, entertained above, that atoms cause an effect by *composing a person who causes that effect*. Similarly, one might claim that whatever I cause by F-ing must be overdetermined, since it will also be caused by you, since you are such that I F. And again, one could say that I overdetermined the Bulls' 1998 victory by being such that Jordan steals the ball from Malone. As already noted (§I above; Chapter 4, §V), I don't think these allegedly overdetermining causes are causes at all. Their obviously riding piggyback on the real causes reveals their impotence. They do not pose a genuine threat of overdetermination.

But there are more troublesome cases we must consider, cases in which it seems that atoms cause—yet do not piggyback upon—what a person causes. Suppose that a human person is tossed through a window. The person, it seems, shatters the window. And surely the person's constituent atoms, working in concert, shatter the window. So the shattering seems to be overdetermined.

First response: Bite the bullet. One could say that when a person is thrown through a window, the window's shattering is overdetermined. One might then add that this is not the sort of *unmotivated* overdetermination resisted in Chapter 3. For in Chapter 3, we had no good reason to believe in the overdetermining culprit—the baseball—in the first place. But we have, for example, first-personal evidence of our own existence and so, by extension, reason to believe other humans exist. This provides positive reason to conclude that, in the case at hand, the window's shattering was overdetermined by a human and some atoms.

As noted in Chapter 3 (§III) and elsewhere, I think we should oppose systematic overdetermination on its own demerits. So I reject this response. Nevertheless, *properly motivated* systematic overdetermination is consistent with my overall ontology and even with my defences of it in this book.

I suppose this 'first response' is an option worth keeping in reserve.

Second response: One could insist that, in general, human organisms cause only what they straightforwardly and directly cause by having mental properties. This implies, one might add, that human organisms *do not* shatter windows by being tossed through them. (Though humans might shatter windows by deciding to shatter them.) So, when a person is tossed through a window, the window is shattered by only the person's atoms acting in concert. Thus there is no overdetermination.

This response maintains that large physical objects can be thrown through closed windows without shattering them. This response initially seems quite bizarre. Nevertheless, while I deny that human organisms cause only what they straightforwardly and directly cause by having mental properties, this response can be made more plausible than one might suspect. But for now, let us set it aside. I'll return to it towards the end of this section.

Third response: The person causes the window to shatter; her atoms cause the window to shatter; but—because the person is causally relevant to whether her atoms shatter the window—that shattering is not overdetermined.

We have already considered cases that are, in some respects, analogous to the picture defended in this response. Suppose, for example, that my *deciding to move my arm* causes the atoms of my arm to cause something (Chapter 4, §V). I and those atoms cause the same effect. Yet it is easy to see that this does not imply any overdetermination.

Some might object that these cases are not at all analogous. In the example involving my deciding to move my arm, my having a mental property straightforwardly and directly causes my atoms to cause something. But, our objectors might insist, in the case at hand—add that the person thrown through the window is in a coma—none of the person's 'decidings' or other

conscious mental properties cause that person's atoms to shatter the window.

In response, perhaps a case like our objectors are insisting upon is not really available. Imagine that you throw me, comatose, through a window. My constituent atoms shatter the window. The causal history of my atoms includes their being thrown by you. But they did not come into the world *ex nihilo* immediately prior to this misadventure. Their causal history also includes my having caused them to move by my conscious mental states such as, for example, my intending some bodily motions.[71] This implies that I cause the atoms—from way up the 'causal chain', by being 'causally upstream' from their causing the shattering—to cause the shattering. And that makes me causally relevant to whether they shatter the window, which precludes my overdetermining that shattering.

There is a second way that I might be causally relevant to whether my atoms shatter a window, even though I am in a coma when they do it. Consider the variety of ways a causal relationship can be instantiated, especially ways that do not involve 'pushing', 'pulling', or other forms of contact. Someone's failing to stop at a traffic light causes an accident. Shaquille O'Neal, by being strategically positioned in the paint, causes the opposing forward to take a jump shot instead of driving in for a lay-up. Someone oversleeps, thus causing them not to show up for guard duty, thus causing the store to be looted during the temporary blackout.

[71] This holds even if, by the time of the window shattering, I am composed of *no* pre-coma atoms. For the atoms that then compose me would have interacted causally with atoms since expelled, which would have interacted causally with other atoms, and so on, until we reach atoms whose behaviour I had directly caused.Perhaps the only actual cases of a human organism's failing to be causally upstream from his or her atoms involve foetuses, assuming that a foetus can exist without ever having been conscious. If, however, some sort of biological anti-reductionism were true (see below and Ch. 4, §VI), even a never-yet-conscious foetus would be upstream from his or her atoms.

Let us focus on this last example. Imagine the security guard causes the looting *by oversleeping.* (He could even cause the looting by slipping into a coma.) Now imagine further that I too oversleep but, being in no special way connected to these events, do not cause the looting. The security guard and I do the same thing—oversleep—but he causes the looting and I do not. Why?[72]

My sketch of an answer is that the security guard has, and I lack, 'potential causal control' over whether the looters loot. The security guard has power—power he does not exercise because of oversleeping—to stop the looting. This is what enables him, by not exercising that power, to cause the looting. And note that he causes the looting *by causing the looters to loot.* (Obviously, the guard is a mere partial cause here.)

The sleeping guard has potential causal control over whether the looters loot. Likewise, even when comatose, I have (because of downward causation; Chapter 4, §V) potential causal control over what my atoms do. The sleeping guard causes the looters to loot. Likewise, I cause my atoms to shatter the window. And so I do not overdetermine the shattering.

A third way to avoid overdetermination in the case at hand requires a substantive assumption. Assume that living organisms, just in virtue of being alive, constantly (and nonredundantly) cause their parts to do things. If so, then just so long as I am alive while thrown through the window, I directly influence the behaviour of my constituent atoms. Living yet comatose, I cause my atoms to cause the shattering of the window much as I would were I intentionally flailing about as I strike the window.

To clarify the point here, note that, on my ontology, there is no single event that is the shattering of a window. Rather, there are the many, many scatterings of the atoms arranged

[72] Suppose the road to causation is so broad that I too cause the looting. This helps. For what matters is that the guard causes the looting (not that I fail to do so); the broader the road, the more room on it for the guard.

windowwise. And whether those scatterings—rather than other, similar scatterings—occur will certainly turn on microscopic differences in my body as I strike the window. This is why my body (i.e. I) would cause 'the shattering'—that is, the many atomic scatterings—if my body, in virtue of being alive, caused my atoms to do things as they shattered the window.

One might object that this line of response helps only if I'm thrown through the window while alive. What if I am thrown through the window when dead? As noted earlier (Chapter 2, §IV; Chapter 5, §II), I think that upon death we cease to exist. Moreover, no corpse mysteriously pops into existence to replace us. Rather, when we die we are 'replaced' by nothing more than atoms arranged corpsewise. Those atoms can shatter a window. But since there is no corpse composed of them, there is no corpse to threaten to overdetermine the shattering.

Atoms arranged corpsewise do, however, suggest a different sort of objection to my claims in this section. For if the day comes when my atoms arranged corpsewise are thrown through a window, there will be a sense in which *I* will cause the window to shatter. After all, even if non-existent, I will have been 'causally upstream' from those atoms. Now this alone is no problem. Suppose you plant a bomb but then cease to exist before it explodes. Though gone, there is a straightforward sense in which you cause the bomb's damage.

The problem is that when I am thrown comatose through a window, and cause it to shatter by being causally upstream from my atoms, I thereby cause the window to shatter *no more directly* than I do when it is shattered by 'my corpse'. In the latter case I cause the shattering *indirectly,* since I don't even exist when it occurs. Yet surely a comatose person thrown through a window ought to *directly* cause it to shatter. To deny this—so this objection concludes—is no better than endorsing the 'bizarre' response that a human can be thrown through a window without causing it to shatter at all.

This is an objection to my claim that, while comatose, I cause the window to shatter by being 'causally upstream' from my window-shattering atoms. And this objection suggests worries about my other two ways of making me causally relevant to my window-shattering atoms, ways involving 'potential causal control' and 'downward causal control in virtue of being alive'. For these ways of avoiding overdetermination imply that I shatter the window *only by causing my atoms to shatter it*. One could thus object that all three of my strategies for avoiding overdetermination (by making me causally relevant to whether my atoms shatter the window) have the same unacceptable implication. They all imply—so this objection goes—that a human thrown through a closed window does not shatter the window in the *direct* way that her atoms shatter it.

My response begins by asking us to suppose that I intentionally punch a window, shattering it (philosophical tough guy). This is to suppose that I use my atoms arranged fistwise, somewhat as I might use atoms arranged baseballwise, to shatter the window. Obviously, when I use a 'baseball' to shatter a window, the 'baseball' and I do not shatter the window in an equally direct fashion. Instead, the 'baseball' shatters the window more directly, more proximately, than do I. Likewise, in the case we are now considering, I cause atoms arranged fistwise to shatter the window more directly than I shatter it. And I think this has a quite general upshot. If *some but not all* of a person's atoms—if, for example, the person's atoms arranged fistwise—cause an effect in a very direct way, there is nothing untoward or counter-intuitive or surprising or objectionable about the claim that the person herself or himself causes that same effect in a less direct way.

Now return to the case at hand. It is not obvious that, when I am thrown through a window, *all* of my constituent atoms cause it to shatter. It could be, instead, that the window is shattered by many or most or nearly all of my atoms working in concert. If so, then there is nothing untoward in saying—as I

do say—that I thereby cause the window to shatter in a less direct way than do the relevant atoms. And so, if the latter is the case, I can block the objection currently at issue here. I think the latter is the case. I think at least one of the atoms in a human body thrown through a window would not contribute causally to the shattering.[73]

These comments offer some support for the second, seemingly 'bizarre' response to the objection at issue in this section. Consider a case in which one's 'punching' the window is unintentional. Suppose that while you are comatose, I use your fist, as if it were a hammer, to shatter the window. It seems quite plausible that, although your atoms arranged fistwise cause the window to shatter, you do not. And, likewise, it is plausible that some (but not all) of my atoms shatter a window when I am tossed comatose through it, but I do not.

But this is not my position. My position is, as already noted, that when tossed comatose through a window I shatter the window—but only comparatively indirectly. I am inclined to say I am a comparatively indirect cause here, as opposed to no cause at all, owing to, in part, the commonsensical sorts of claims rallied at the start of Chapter 4 in support of our being organisms. I can be seen. But if I can be seen, then I myself—not just some of my atoms—must be somehow causally involved in producing the relevant visual sensations. And it seems odd to think that I am so involved when awake and intentionally moving about, but not thus involved when comatose. A blow to the head might render me unconscious, but not invisible.

At any rate, I think we can easily handle the sort of challenges posed thus far with respect to what one causes that is

[73] One might be tempted to 'help' by adding that, when a human is thrown through a window, only her *external surface* (i.e. her atoms arranged external-surfacewise) shatters the window. This temptation should be resisted. If only the body's surface had struck the window—suppose it had been 'shaved' off the body—the window would not have shattered. The force with which an object strikes (or objects strike) something is in part a function of its (their) mass; the mass of a human's 'surface' is far less than that of the human.

also caused by some, but not all, of one's atoms. But there might be some things that I seem to cause that also seem to be caused by *all* of my atoms—every single one—working in concert. It is difficult, however, to come up with an obvious example of such a thing. But if faced with such an example, I'd say that it is an example in which I cause the effect comparatively indirectly, by causing my atoms to cause it (by my being causally upstream, etc.). After all, given that I only indirectly cause a window to shatter when I am thrown comatose through it, it isn't so surprising that I would indirectly cause what would be caused by all of my atoms working in concert.

In Chapter 3, I said that, if baseballs existed, they would shatter windows. Given the above, one might say that, if baseballs existed, then some (most) of their constituent atoms would, working in concert, shatter windows—but baseballs themselves would not. Thus, one might claim, it does not follow from the baseball's failing to shatter the window that it does not really exist.

This claim—if meant to be a criticism of my arguments—is beside the point. I never suggested that the shattering of windows is the sine qua non of baseball existence. The point is that—as explicitly noted when considering the moral of the Overdetermination Argument—overdetermination-based considerations give us a reason to say that, if baseballs existed, then they would cause *absolutely nothing*. Yet if they existed, they would surely cause *something*. And so we ought to deny that baseballs exist.

Indeed, I think that, if only *some* of a baseball's atoms—and so not the baseball, even if it existed—would shatter a window, we have further support for eliminativism. For if baseballs wouldn't cause windows to shatter, what would they cause? In so far as coming up with a clear example is difficult, making the claim that baseballs cause things even prima facie plausible is difficult. And this alone—setting aside the arguments of Chapters 2 and 3—undermines the prima facie plausibility of

the claim that baseballs exist. It is, on the other hand, quite easy to come up with clear examples of what we humans cause. For we cause whatever is caused by (human) mental causation. This is more evidence that it is not arbitrary to eliminate baseballs (and statues) but not human organisms.

III. The 'Bottom-Up' Threat to Free Will

Some philosophers think that human freedom and 'determinism' are incompatible. These 'incompatibilists' think that, if everything one does is entailed by what the distant past was like conjoined with the nature of the laws of nature, then one cannot act freely. The principal motivation for incompatibilism—or at least the motivation that has received the lion's share of attention over the past thirty years—is the 'no choice' argument (cf. Ginet 1966; Wiggins 1973; van Inwagen 1983. Here is a version of that argument:

(i) Humans have no choice about the following truth: every action a human performs is entailed by what the distant past was like and the nature of the laws of nature.
(ii) Humans have no choice about what the distant past was like or the nature of the laws of nature.

Therefore,(iii)

Humans have no choice about what actions they perform.

This argument purports to show that determinism—the thesis that premiss (i) alleges is a truth about which we have no choice—leads to lack of choice, and so to lack of free will, with respect to all human actions. Some incompatibilists accept (iii) and conclude that we have no free will; most reject premise (i), concluding that determinism is false.

Not everyone agrees that this argument forces us to choose between determinism and free will. For some respond to this argument by denying premiss (ii) instead. They think that there is *some* sense in which we 'have a choice about' the past or the laws of nature. After all, they argue, we can do things such that, were we to do them, either the past or the laws of nature would have been different (Lewis 1981; Horgan 1985).

But the most extensive literature on this sort of argument has to do with whether it is valid. And controversy over its validity just is controversy over various closure principles for 'having no choice'. For example, and most simply, if 'having no choice' is closed under entailment, the argument is valid. The question, then, is whether *any* of the closure principles that render the argument valid are true.[74]

I think the 'no choice' argument is valid and incompatibilism true (and determinism false). But I shall not defend this here. Rather, I want to show only that the very sort of reasoning that leads many to incompatibilism should lead them to a further claim about freedom.

Consider the following argument:

(I) Humans have no choice about the following truth: every action a human agent performs supervenes on what that agent's constituent atoms do or are like.
(II) Humans have no choice about what their constituent atoms do or are like.
Therefore,
(III) Humans have no choice about what actions they perform.

[74] Van Inwagen's 'Beta Principle' (1983: 94) has received the most scrutiny: N *p* and N(if *p* then *q*) entails N *q,* where 'N *p* ' means '*p* and no one has, ever had, or will have a choice about *p* '. See discussions in Widerker (1987), McKay and Johnson (1996), Kane (1998, ch. 4), Crisp and Warfield (2000), Finch and Warfield (1998), and O'Connor (2000, ch. 1).

We should read the above argument so that premiss (II) does not beg the question. Premise (II) would beg the question if what one's constituent atoms 'are like' included their being such that one performs various actions. For example, that I have no choice about my atoms' being such that I perform action A presupposes that I have no choice about performing action A. (A similar point holds, of course, for how we read the classic 'no choice' argument. In reading it, we allow neither what the distant past 'was like' nor the 'nature' of the laws of nature to include my being such that I perform action A.)

We shall understand claims about what one's atoms 'are like' in such a way that (II) does not presuppose (III). Not only does this keep the argument from begging the question, it also keeps the supervenience affirmed in (I) from being trivial. For if what one's atoms 'are like' included their being such that the human they compose performs action A, a human's performing action A would trivially supervene on what his or her atoms are like.

So the above argument does not beg the question. Thus it is not—in virtue of begging the question—trivially valid. Is the argument valid at all? This is a disputed question. For it is valid if and only if the classic 'no choice' argument is valid. This is because any closure principle for 'having no choice' that renders one valid will do so for the other. As noted, many think the classic 'no choice' argument is invalid. They will claim that the above argument is likewise invalid. But every *incompatibilist* ought to think the above argument is valid. For, surely, every incompatibilist ought to believe that the original 'no choice' argument is valid.

Let me put the same point a different way. Call the conjunction of premisses (I) and (II) 'bottom-up metaphysics of human actions'; or, for short, 'bottom-up metaphysics'. The above argument, given that it is valid if and only if the classic 'no choice' argument is valid, shows that *bottom-up metaphysics is as much of a threat to human free will as is determinism*. (More

carefully, it shows that bottom-up metaphysics is as much of a threat to free will as the conjunction of premisses (i) and (ii) of the classic argument, which premisses entail determinism and then some.) Now that could mean that human free will is threatened by neither bottom-up metaphysics nor determinism. Or it could mean that bottom-up metaphysics and determinism each rule out free will entirely. It all depends on whether a suitable closure principle is true.

Something like a 'bottom-up threat' to human free will is occasionally tossed into the pot along with determinism (see the Cicero quote below). But that threat is rarely carefully developed and explicitly distinguished from the more familiar challenge posed by determinism (but see Cover and O'Leary-Hawthorne 1996). And so, compared to the threat of determinism, the bottom-up threat is relatively neglected. This neglect is unfortunate. After all, bottom-up metaphysics is widely held. Indeed, I think this metaphysics is more widely endorsed than is determinism, because the latter is presumed to be inconsistent with actual quantum indeterminacy.

Mention of quantum indeterminacy highlights another reason the above argument is important. Some have thought such indeterminacy—or something for present purposes just like it—was relevant to securing human freedom. For example, as Cicero tells it:

> [Epicurean philosophers postulated the chance 'swerve' of atoms to allay their] fear lest, if the atoms were always carried along by . . . natural and necessary force[s] . . . we should have no freedom whatever, since the movement of the mind was controlled by the movement of the atoms (quoted in Kane 1998: 6).

It may be old news that quantum indeterminacy and atomic swerve are not sufficient for free will. Yet the Epicurean suspicion that one or the other is necessary (given incompatibilism) may still be quite common. This suspicion is misguided. For if determinism precludes freedom, then so does bottom-up

metaphysics. So, given incompatibilism, human freedom requires (at least) one of the following two things. A person has some choice about what her atoms do or are like (the denial of (II)). Some of a person's actions fail to be fixed, one way or another, by atomic behaviour or features (the denial of (I)). If we have either, quantum indeterminacy (atomic swerve) is not needed for freedom. If we have neither, quantum indeterminacy (atomic swerve) won't help. As a result, quantum indeterminacy (atomic swerve) turns out to be *irrelevant* to human freedom.

Every incompatibilist ought to grant that, if we ever act freely, one of the two premisses of the above argument is false. And we have already been given reasons to oppose these premisses, reasons that have nothing to do with free will. If a human's having a conscious mental property counts as his or her performing an 'action', then Chapter 4's (§II) denial of C undermines (I). But I won't dwell on this, since I don't want to quibble about what counts as an 'action'. Besides, whatever we say about (I), we can see that (II) is false. For Chapter 4 (§V) showed us that human persons have downward causal control over their constituent atoms. And surely downward causal control of this sort is *sufficient* for having a choice about what one's atoms do or are like.

Moreover, downward causal control is *necessary* for having such a choice, at least given the sense of 'choice' at play in the arguments we've been discussing. To see why I say this, recall that some object to (ii) above by maintaining that there is a *sense* in which we have a choice about the laws of nature or the past. This objection leaves many unmoved. Why? Because—so I suggest—the sense of 'choice' involved in the objection is weaker than the sense involved in the argument. The sort of choice addressed in the argument has a causal component, but—I say—not the sort of choice relied upon in the standard objections to premiss (ii). Given its causal component, I conclude that we have the relevant sort of choice about what our

atoms do or are like only if we have downward causal control over them.

So (II) is false if and only if we have downward causal control over our constituent atoms. On the assumption that we are human organisms, I have argued that we exercise downward causation. (II) is false.[75]

I say that the downward causal control we exercise over our atoms makes room for our having free will. And, as we saw in the previous section, that same downward causal control undermines the Micro Exclusion Argument for mental epiphenomenalism. I think free will requires mental causation. So I think it bodes well for my metaphysics that its defence of free will turns on the same fact about humans as does its defence of mental causation.

At the end of § I, I noted that we could invert the argument of that section in order to defend the claim that humans have downward causal control. We can do something similar here. We could argue that, given incompatibilism, human freedom implies that either (I) or (II) is false. But we sometimes act freely. Therefore, we should deny (I) or (II). If (I) is false, then so is the bottom-up metaphysics of human persons according to which all we do supervenes on what our parts are like. If (II)

[75] Here is one moral of the above discussion. Incompatibilism is as open to those who believe that we are physical objects as it is to, for example, substance dualists. Our being organisms is no reason to reject incompatibilism conjoined with the claim that we are free. Maybe this is surprising. Timothy O'Connor (1995: 179) says 'many philosophers who discuss the agency theory [understood by O'Connor as the only incompatibilist avenue to free will] seem to simply *assume* that its adherents are dualists'; in this regard he cites Honderich (1988) and Levison (1978). And Cover and O'Leary-Hawthorne (1996) charge that non-dualists can resist the bottom-up threat only in an *ad hoc* manner, a charge that my argument in this section shows to be mistaken.It is worth pointing out that the (arguably) leading defender of incompatibilism and free will—Peter van Inwagen—is also the (arguably) leading defender of the claim that we are human organisms (see van Inwagen 1983, 1990). Of course, one could argue that R. Kane is the leading defender of incompatibilism who believes in free will; yet Kane (1998: 118–19) explicitly denies that incompatibilism should push one towards substance dualism. And O'Connor argues for incompatibilist freedom on the assumption that we are physical entities (2000, ch. 6).

is false, we have downward causal control over our parts, and the bottom-up metaphysics of humans is false. For many of us, the 'top-down' metaphysics of human persons defended in Chapter 4—the metaphysics that saves us from the Overdetermination Argument—is overdetermined.

IV. Conclusion

The arguments of Chapters 2 to 5 show that we should eliminate baseballs and statues but not humans. But those arguments aside, one might query, doesn't it *seem* like statues and human organisms should be in the same metaphysical boat? Isn't it intuitively arbitrary to eliminate composite statues but not composite persons?

No. As I emphasized at the close of §VI in Chapter 4, the ontology here can be well motivated by favouring objects with non-redundant causal powers and thus downward causal control over their proper parts. This chapter—relying as it has, in each section, on either our non-redundant powers or our resultant downward causal control—casts further light on the differences, relevant to ontology, between alleged statues and actual humans.

7 Belief and Practice

SOMEONE who tries to purchase statues is—in an important sense—like someone who tries to hunt unicorns. Both are doomed to failure. For there are neither statues in the galleries nor unicorns in the woods. Someone who tries to acquire a statue is—in an important sense—not remotely like someone who tries to bag a unicorn. One needs money, the other therapy. In this chapter I shall explore how, given the truth of eliminativism, belief in statues (and in the other inanimate macrophysical objects of folk ontology) resembles, and how it differs from, belief in objects like unicorns.

I. False Folk Beliefs

In Chapter 1 (§III) I considered the claim that folk uses of sentences like 'there are statues' or 'statues exist' mean only that there are some things arranged statuewise. This claim was part of a challenge to the very coherence of eliminativism. But, I argued, this claim about meaning is false. And so for this reason, among others, that challenge failed.

I shall here reconsider this claim about meaning. But I am not now concerned with challenges to eliminativism's coherence or even to its truth. I am instead concerned with Peter

van Inwagen's attempt to render eliminativism consistent with ordinary beliefs. I am concerned with his project of endorsing eliminativism while also maintaining that folk claims about statues—and the beliefs such claims reflect—are true. The key to van Inwagen's reconciliation is his philosophy of language. According to that philosophy of language, when the folk say 'statues exist', they report only that there are some things arranged statuewise (1990, §§10–11).[76] Thus we return to the kind of claim about meaning I opposed in Chapter 1.

I still oppose such claims. One reason to oppose them stems from eliminativism's undeniably striking and surprising nature. For—ask yourself—*why* is eliminativism striking and surprising? It cannot be because of its revisionary practical or empirical consequences; it has no such consequences (Chapter 1, §II; and see §III below). Instead, eliminativism is striking and surprising simply because—and this is the obvious answer—it contradicts what nearly all of us believe.

Van Inwagen must reject this obvious answer. Indeed, I think that van Inwagen must say that eliminativism contradicts *only* some philosophers *qua* philosophers. But the claim that eliminativism is consistent with what all people believe—except for a comparative handful of professional ontologists, and even them only while they are working—does not do justice to its striking and surprising nature.

Moreover, suppose that, when the folk say 'statues exist', they commit themselves only to the claim that there are things arranged statuewise. This implies that, in the folk lexicon, 'statues exist' does not mean that something exists that is a statue. In the parlance of Chapter 1 (§III), this implies that, in ordinary uses of 'statues exist', 'exist' is being used 'deviantly'.

[76] Van Inwagen does not mean by locutions like 'arranged statuewise' exactly what I do. My account of such locutions is in terms of counterpossibles (Ch. 1, §I); his is not (van Inwagen 1990: 109). Because our differences here are irrelevant to the argument of this section, I shall ignore them.

So van Inwagen's position implies that the folk's assertion of 'statues exist' should not be taken *completely literally and straightforwardly*. Those who follow van Inwagen here must say that the content of 'statues exist' in ordinary discourse is not right on the surface. Or they must say that 'statues exist' doesn't wear its logical structure on its sleeve. Or they must say something else along these lines. And van Inwagen *does* say something along these lines. The best way to present both van Inwagen's position, and his principal defence of that position, is to quote the following passage from his *Material Beings*:

> When I speak the words 'the sun moved behind the elms', I am reporting a fact. I am reporting a real alteration in the relations of external objects. Perhaps the words I use constitute what is in some sense a *misleading description* of this fact, but they do at least get one thing literally right: Taken literally, they report an alteration in the spatial disposition of external objects and an alteration in the spatial disposition of external objects really does occur and is the basis for the report. Thus, 'The sun moved behind the elms' is not, even from the point of view of the most fanatical astronomical literalist, a report of a nonexistent, fabricated, or imaginary event; it is not like, say, 'The sun moved rapidly back and forth across the sky'. It may describe an actual event in a *misleading or loose or even wrong way,* but the event it describes or misdescribes is there to be described or misdescribed. Something similar may be said about 'There are two very valuable chairs in the next room'. (1990: 101–2; emphasis added)

Van Inwagen says that in ordinary contexts 'there are two very valuable chairs in the next room' expresses a proposition about things arranged chairwise, but not chairs. And I presume that he would say something similar about sentences like 'there are chairs' or 'chairs exist'. After all, even if 'chairs exist' often goes without saying, whatever that sentence would express in ordinary discourse rarely goes without believing. So in what follows I'll consider the alleged analogy between folk uses of 'chairs exist' and ordinary uses of 'the sun moved behind the elms'.

This analogy does double duty for van Inwagen. It is both a statement of, and the chief support for, his position on folk uses of 'chairs exist'. That is, first, van Inwagen's position can be summarized by the claim that ordinary uses of 'chairs exist' are relevantly like ordinary uses of 'the sun moved behind the elms'. And, second, his primary defence of that position seems to rely on the hope that, once we put it in terms of this analogy (or others like it[77]), we will find it plausible.

'Absolute Ptolemaists' believe the earth is absolutely fixed and the sun moves round it. Imagine an Absolute Ptolemaist who uses 'the sun moved behind the elms' to express literally and directly and non-figuratively a claim about the sun's movement with respect to absolutely stationary elms. Her use of that sentence expresses a *false* proposition. Yet it is obvious that 'the sun moved behind the elms' in ordinary twenty-firstcentury discourse does not express that same false proposition. This is obvious because we twenty-first-century discoursers do not believe that the sun moves relative to absolutely stationary elms. When we use 'the sun moved behind the elms' we report something, but not that the sun thus moved. So I suggest we grant van Inwagen his claim that that sentence, nowadays at least, ordinarily expresses a fact in a 'misleading or loose or even wrong way'.[78]

[77] Van Inwagen invokes three more analogies (one in the form of a fable). For criticism of these, see Mackie (1993), Michael and O'Leary-Hawthorne (1996, §II), and Noonan (1999).

[78] Some might object that even taken literally and straightforwardly, 'the sun moved behind the elms' is true. The basis of this objection is that the most literal sense of 'motion' is relative; thus, since the sun moves relative to the elms—and so moves in the literal sense—our target sentence is, taken literally and straightforwardly, true (cf. Michael and O'Leary-Hawthorne 1996). This objection is probably mistaken, since acceleration (and so circular motion) is not—at least not in Special Relativity—appropriately relative. And such details aside, surely there is some sense in which, with respect to what moves around what, Ptolemy got it wrong and Copernicus got it right! At any rate, for the sake of argument, I shall grant van Inwagen's assumption that 'the sun moved behind the elms', if it were interpreted in the most literally wooden manner, would express a falsehood.

The following dialogue illustrates the point here:

A. The sun moved behind the elms.
B. Are you trapped in a medieval cosmology? Do you affirm 'the sun moved behind the elms' in the most literal and straightforward fashion possible?
A. Of course I don't. You ask funny questions.

But now imagine the following:

A. There are chairs over there.
B. Do you mean to say that there are some things over there such that they are chairs? Do you affirm 'chairs exist' in the most literal and straightforward fashion possible?
A. Of course I do. You ask funny questions.

These exchanges highlight an important point. Ordinary speakers are, I believe, happy to concede that 'the sun moved behind the elms' is in a certain sense a misleading or loose or even wrong way of characterizing a claim that is not committed to the literal movement of the sun. Ordinary speakers are not, I believe, happy to concede that 'chairs exist' is likewise a misleading or loose or even wrong way of characterizing a claim that is not committed to the literal existence of a chair.

The literal and straightforward meaning of claims like 'chairs exists' or 'there are statues' is indeed the ordinary, folk meaning. That seemed like the right thing to say in Chapter 1 (§III). And when we consider the most prominent proeliminativist challenge to this claim—van Inwagen's example of the sun moving behind the elms—we find, instead of a reason to reject this claim, further support for it. For van Inwagen has provided a nice illustration of the following point. We must choose between interpreting a speaker as speaking in a literal and straightforward way and interpreting her as speaking in a somewhat misleading or loose or even wrong way. Claims like 'the sun moved behind the elms' ordinarily express truths in part because they ordinarily express

those truths loosely or somewhat misleadingly. Since 'chairs exist' is ordinarily not a misleading or loose or wrong way to say something, it must ordinarily be interpreted literally and straightforwardly. Thus interpreted, the eliminativist insists, it is false.

There is another reason to hold that ordinary uses of 'chairs exist' are properly interpreted straightforwardly and literally. The first step in presenting this reason is to make a conjecture. I surmise that if van Inwagen became convinced that folk ontology were true, he would conclude that ordinary uses of 'chairs exist' should be taken straightforwardly and literally. For his motivation for finding a way to take 'chairs exist' non-straightforwardly or non-literally seems to be the desire to make ordinary uses of that sentence express truths. Surely if such sentences were true taken straightforwardly and literally, there would be neither the need nor the will to take them some other way.

Add to the point just noted that the folk are humans. Thus whenever the folk say or think *anything,* humans exist. So whenever the folk say 'humans exist', humans exist. Now 'humans exist' *seems* to mean that humans exist; that seems to be the most natural interpretation. And if that is how we interpret folk uses of it, it is always true. Thus there is no compelling motivation for interpreting 'humans exist' any other way. (The only motivation might come from wanting 'humans exist' to follow the precedent allegedly set by 'chairs exist'.)

That's my argument for the claim that 'humans exist', in the folk idiolect, means that humans exist. And something else makes me want to endorse this claim. If our ordinary beliefs about 'humans' are merely beliefs about things arranged humanwise, then those beliefs do not sanction the positive part of my ontology: the existence of us human organisms. And so I could not maintain, as I have been all along, that the belief in humans is initially presumed innocent on the grounds that it is part of common sense.

When the folk say 'humans exist', they say that humans exist. But then something similar holds for when they say 'chairs exist'. For it is not plausible that folk uses of 'humans exist' report the existence of humans but folk uses of 'chairs exist' do not report the existence of chairs. Thus we have a third sort of reason to interpret ordinary uses of 'chairs exist' literally and straightforwardly and so to reject any philosophy of language that tells us to interpret them otherwise. (The first reason, recall, dealt with doing justice to eliminativism's striking nature; the second with considerations surrounding the sun and the elms.)

Folk uses of 'chairs exist' should be taken literally and straightforwardly. All along, I have been assuming that this implies—given eliminativism—that such uses of 'chairs exist' express falsehoods. And so they do. But one might object to this assumption. The objection I have in mind begins by supposing that there is more than one equally literal and equally straightforward meaning of 'exist'. Some, like Quine (1953), Lewis (1986*a*: 212), and van Inwagen (1998: 236–7), have directly opposed this supposition. But it has prominent defenders including Gilbert Ryle and (I hazard to suggest) Wittgenstein (1961: 4.1272). For example, Ryle says:

> It is perfectly proper to say, in one logical tone of voice, that there exist minds and to say, in another logical tone of voice, that there exist bodies. But these expressions do not indicate two different species of existence, for 'existence' is not a generic word like 'coloured' or 'sexed'. They indicate two different senses of 'exist', somewhat as 'rising' has different senses in 'the tide is rising', 'hopes are rising', and 'the average age of death is rising'. A man would be thought to be making a poor joke who said that three things are now rising, namely the tide, hopes and the average age of death. It would be just as good or bad a joke to say that there exist prime numbers and Wednesdays and public opinions and navies; or that there exist both minds and bodies. (1949, 23)

Ryle's position would lose its punch if he added that, say, only prime numbers *really* exist, 'Wednesdays exist' (and the rest)

being true only when 'exist' is used deviantly. But of course Ryle does not add this. Indeed, he would surely deny it. Ryle's position is not only that we use 'exist' in different ways when applied to things of different 'categories', but that those different ways are all *equally straightforward and literal*.[79]

I reject Ryle's position. And I am tempted to peter out by saying I have no idea what these multiple senses of 'exist' could be and to leave it at that.[80] But perhaps we can say a bit more. One objection to Ryle begins by noting that we can, without fixing which of the allegedly many senses of 'exist' applies, truly claim that *something* exists (cf. Benardete 1989: 46–7). Relatedly, one could object that Ryle is committed to claims like 'for every *x*, if *x* exists-in-the-sense-of-'exist'-predicated-of-prime-numbers, then *x* does not exist-in-the-sense-of-'exist'-predicated-of-Wednesdays'. But a claim like this, a claim about *every* (existing) thing, seems to presuppose a wholly unrestricted and univocal kind of existence.

I find these comments telling. But I shall not develop them further. For my primary aim is not to establish, to everyone's satisfaction, that the Rylean view is false. Rather, it is only to establish that no eliminativist should find in it a reason to think folk beliefs are true. So note that eliminativists should reject the following:

> 'Exist' has one literal and straightforward meaning in folk uses of 'chairs exist', which uses express a true proposition and a true belief. 'Exist' has a distinct, but equally literal and straightforward meaning, when the eliminativist speaks truly by saying 'chairs do not exist' and 'humans exist'.

[79] Ryle's disavowal of different species of existence brings to mind Aristotle's claim that 'being is not a genus'. But Aristotle privileges the existence enjoyed by substances (like human organisms) over the existence enjoyed by non-substances (like statues). And so it is not clear whether he holds that 'exist' is applied in a completely straightforward and literal way to non-substances. See Barnes (1995: 72–7) for an introduction to Aristotle's views on this point.

[80] peter out, v. to claim not to understand one's opponent; hence, derivatively, as of an argument, to end because one or more of the disputants peters out. 'The session petered out when the speaker quined herself.'

The position just stated promises to make 'eliminativism' consistent with folk beliefs and so, by the reckoning of some, more palatable. But, we have seen, such consistency is itself problematic. For recall that if eliminativism were consistent with folk beliefs, it would be neither surprising nor striking; but it is surprising and striking; so we should reject the above position.

More importantly, if that position were correct, eliminativism would not be an interesting philosophical thesis. It would be, instead, a silly fixation with one out of many equally weighty kinds of existence, a fixation with the kind of existence humans have but chairs lack. This is no 'eliminativism' of chairs worthy of the name. So no eliminativist should endorse it.

I think there is but a single literal and straightforward sense of 'exist'. Yet someone who disagreed with me on this could defend something that legitimately counts as eliminativism. For she could defend the claim that chairs and statues and baseballs exist in *none* of the literal and straightforward (and supposedly many) senses of 'exist'. Her defence might start by noting that the arguments of the previous chapters for the falsity of 'statues exist' never relied on just one literal sense of 'exists', never required us to set aside the other literal senses.

The folk speak literally and straightforwardly when they say 'chairs exist'. And the eliminativist should hold that 'chairs exist', on any literal and straightforward interpretation, expresses a falsehood. Thus I conclude that all eliminativists should agree that ordinary uses of 'chairs exist' (and 'baseballs exist' and 'statues exist') express falsehoods. Likewise, eliminativists should agree that when the folk believe something which they would ordinarily express with 'chairs exist' (or 'baseballs exist' or 'statues exist'), they believe falsely. It is easy to see why eliminativism is striking and surprising.

II. False Folk Beliefs Are Nearly as Good as True: Justification

As van Inwagen rightly notes (1990: 101–2), the eliminativist who says that ordinary ontological beliefs are false must award such beliefs some sort of 'alethic commendation'. So I must explain the sense in which the belief ordinarily expressed by 'statues exist' is *better than* the belief expressed by 'unicorns exist', given that both are false. (For ease of exposition, below I shall make claims about sentences like 'statues exist' and 'unicorns exist'. Such claims, as should be evident from their context, will often be shorthand for claims about the propositions or beliefs expressed by those sentences.)

I commend false folk-ontological beliefs for being—here I introduce a technical expression—'nearly as good as true'. Any folk-ontological claim of the form 'F exists' is *nearly as good as true* if and only if (i) 'F exists' is false and (ii) there are things arranged F-wise. So, for example, 'the statue *David* exists' is nearly as good as true because (it is false and) there are some things arranged Davidwise.[81]

'*David* exists' is nearly as good as true. 'Unicorns exist' is not. That's a difference between the two. But, if I am to provide an alethic commendation for '*David* exists', I must explain how that difference on the part of '*David* exists' is commendable.

People who believe in unicorns are few and far between. And those few are generally unjustified. On the other hand,

[81] Theodore Sider, in offering 'words of comfort' to 'lovers of desert landscapes everywhere' (1999*a*: 325), gives an account of 'quasi-truth' and 'underlying truths'. Being nearly as good as true is sufficient for being quasi-true in Sider's sense. And my nearly-as-good-as-truth makers are 'underlying truths', given Sider's definition. Thus what I say here fits into at least one model that is meant to handle, in an entirely general fashion, the way in which eliminativists–nominalists about Fs can agree that 'F exists' is, if not true, at least somehow commendable.

people who believe in statues are legion. And they are generally justified in so believing. Given the truth of eliminativism, we might ask *why* the belief in statues is more common, and more commonly justified, than the belief in unicorns.

The answer is that statue beliefs are nearly as good as true. For, so I claim here, *atoms arranged statuewise* often play a key role in producing, and grounding the justification of, the belief that statues exist. In general, a false belief's being nearly as good as true explains how *reasonable* people come to hold it. And, relatedly, its being nearly as good as true can ground its justification. Because the belief that unicorns exist is not nearly as good as true (i.e. because there are no things arranged unicornwise), there is no similar explanation of its production or similar reason to think it is justified.

Moreover, in Chapters 1 (§II) and 3 (§III) I defended the thesis that whether atoms arranged statuewise compose a statue was not a straightforwardly empirical question. If the question of whether folk-ontological claims are true or only nearly as good as true is not straightforwardly empirical, folk-ontological claims cannot be condemned by straightforwardly empirical facts. And failing to be condemned by straightforwardly empirical facts is commendable. Even if such praise is faint, it's praise that—worries about conclusive empirical evidence for negative claims notwithstanding—the belief in unicorns fails to garner.

Nevertheless, the fact that the straightforwardly empirical evidence is consistent with both folk ontology and eliminativism could motivate an objection. Recall that in Chapter 3 (§III) I argued that once we accept this fact, one who has only sensory evidence for the existence of statues should not believe that statues exist (i.e. should either withhold belief or positively disbelieve). And so one might object that statue beliefs, in spite of being nearly as good as true, are not justified.

In reply, recall that in Chapter 3 I maintained that beliefs like 'statues exist' are initially justified. I argued, however, that

their justification is undermined for those of us familiar with the arguments of §III of Chapter 3. So false folk-ontological beliefs are justified (in some central sense of justification) for those unaware of how that justification can be undermined. This is not the case for belief in unicorns. Moreover, false folk-ontological beliefs (unlike unicorn beliefs) are prima facie justified even for those of us aware of how that justification can be undermined. They are not, of course, ultima facie justified for us.

False folk-ontological beliefs are justified, in various ways for various people, in virtue of being nearly as good true. But they are never 'warranted'. (Warrant is that, whatever it is, that makes the difference between mere true belief and knowledge.) I say false folk-ontological beliefs are never warranted because, as I have argued elsewhere (1995*b,* 1997*b*), no belief can be both false and warranted. (Obviously, claiming that false folk beliefs are not warranted does not stack the deck in my favour. Quite the contrary. For if they were warranted, they would have another leg up on the presumably unwarranted belief that unicorns exist.)

But false beliefs, if otherwise commendable, can have something to do with the warranting of other beliefs. Imagine a society of Hellenic fundamentalists. Being Hellenic, they say things like 'Apollo's chariot moved behind the elms'; being fundamentalists, they thereby mean that Apollo's chariot moved behind the elms. Such sentences express unwarranted falsehoods, even when they correspond to the sun and elms' moving relative to each other in the way that we would (misleadingly, loosely, wrongly!) describe by saying 'the sun moved behind the elms'. Now suppose a first member of our imagined society says to a second 'I shall be drinking tea when Apollo's chariot moves behind the elms'. I think that, if other conditions are right, the second can—upon seeing the sun move behind the elms—come to know (and so become warranted in believing) that the first is drinking tea.

Similarly, false folk-ontological beliefs have something to do with the warranting of other beliefs. There is a restaurant in Texas where, to indicate desire for another plate of the all-you-can-eat fare, one raises 'a table-top Mexican flag'. Suppose your server concludes that you want more food from her belief that you have hoisted your flag. There are no flags. Her belief is false, thus unwarranted. Yet, assuming other conditions are right, it plays a role in her coming to know (and so in coming to be warranted in believing) that you want more food.

Again, suppose that, knowing my address, you locate my atoms arranged housewise. You then conclude that 'Merricks lives here' on the basis of your nearly as good as true belief that my house is in front of you. There are no houses. Your house belief is false. And so unwarranted. Nevertheless, assuming other conditions are right, that belief plays a key role in your knowing (and so in your being warranted in believing) 'Merricks lives here'.

How *exactly* nearly as good as true beliefs are involved in the warranting of other beliefs—the epistemological 'diagnosis' of such cases—is not obvious. I'll suggest two ways this could go in the 'Merricks lives here' case, just to illustrate that there is more than one option. First, perhaps what warrants your belief about where I live is simply its being appropriately based on your (unwarranted but) nearly as good as true belief about a house. Or, second, perhaps your belief 'Merricks lives here' is warranted by a true, warranted, and tacit belief that always accompanies your house belief; perhaps your true belief about where I live is warranted because it is appropriately supported by your (tacit) true and warranted belief about some things arranged housewise.

It does not matter which of these two diagnoses—or whether some third diagnosis—is correct. What matters is that, as the examples above show, nearly as good as true beliefs are intimately linked to the warranting of other beliefs. This is not so for merely false beliefs, at least not those like 'unicorns

exist' which lack some obvious commendation. Nearly as good as true beliefs thereby have an advantage over the merely false.

This is a good place to respond to an objection that may have occurred to some while reading earlier parts of this book. The justification and warrant for our belief in microscopica—and even in the 'bottom-up' metaphysics applied to artefacts—depend upon beliefs about laboratory equipment, such as beliefs about the position of a needle on a dial. Thus one might charge that my overall ontology is epistemically self-defeating. One might charge that since, as I maintain, there is no laboratory equipment, there is no way of knowing about the microscopica that I say are arranged in various ways, such as statuewise. In response, just so long as our 'laboratory equipment beliefs' are nearly as good as true, we can make use of those beliefs (or the true and warranted tacit beliefs which are always linked to them) in coming to know about atoms or other microscopic entities.

III. False Folk Beliefs Are Nearly as Good as True: Practice

'Statues exist' is false but nearly as good as true. 'Unicorns exist' is merely false. The nearly as good as true, as I argued in the preceding section, are better than the merely false with respect to a cluster of epistemic norms. As I shall argue in this section, the nearly as good as true are also better than the merely false with respect to certain practical issues.

Eliminativism, because of its commitment to nearly as-good-as-truth makers for false folk ontological claims, makes little practical difference. For all practical purposes, it doesn't matter whether the batter swings at a baseball or at atoms arranged baseballwise. Nor does it make a practical difference whether we say (and believe) that the batter swings at a

baseball or at atoms arranged baseballwise. In this spirit, I shall say that—given eliminativism—claims about alleged inanimate macrophysical objects are 'true for practical purposes' just in case they are nearly as good as true. (On the other hand, given folk ontology, such claims are 'true for practical purposes' just in case they are true. More on this below.)

Eliminativism can accommodate our practices. This point was made in Chapter 1 (§II) and it should be fairly obvious. Rather than multiply examples in support of this obvious point, I want to defend something less obvious: in many cases eliminativism *better* accommodates our practices than does folk ontology. In defending this claim, I redeem promissory notes issued in Chapter 1 (§II) and in Chapter 3 (§III).

My argument begins with the claim that, for practical purposes, we often assume that the identity over time of inanimate macrophysical objects can be somewhat conventional. For example, we might, for practical purposes, leave it up to the courts to 'decide' a case of statue identity over time that—prior to any judicial decree—is in some sense borderline. (Just to keep things manageable, I shall focus on decisions by 'the courts'. But, of course, the courts play only a minor role in fixing which conventions we adopt. And so they play only a minor role in fixing by convention which claims about identity over time are 'true for practical purposes'.)

Since there are no statues, *all* allegations that a statue enjoys identity over time are false. Yet some such claims are otherwise commendable because they are nearly as good as true. 'S at *t* is the same statue as S* at *t**' is nearly as good as true just so long as (it is false and) the atoms at *t* are arranged *same-statuewise* as the atoms at *t**. This invites a definition of 'same-statuewise'. Here it is:

> Atoms at *t* are arranged *same-statuewise* as atoms at *t** if and only if (i) the atoms at *t* are arranged statuewise; (ii) the atoms at *t** are arranged statuewise; and (iii) *if* there were

> persisting statues, *then* the atoms arranged statuewise at *t* would compose the same statue as the atoms arranged statuewise at *t**.

It is somewhat a matter of convention whether atoms at *t* are arranged *same-statuewise* as atoms at *t**. This is, in part, due to the counterpossible that is clause (iii) of the above account.[82] For it is somewhat a matter of convention whether this counterpossible is (non-vacuously) true. As we shall see, this is due to the fact that part of what grounds this counterpossible is what the folk mean by 'same statue'. (The features and relations of the atoms involved also partially ground that counterpossible.)

Imagine a case of statue identity described by a folk ontologist as follows:

> It was vague, indeterminate, whether my statue was identical with the one owned long ago by the royal family. For a variety of reasons, the whole matter ended up in court. The courts decided that it was indeed the same statue. So it's no longer vague. It is determinately the same statue.

Here is how I interpret the incident just described. The courts' decree creates a context in which it becomes appropriate to invoke one (or more) of many possible precisifications of 'same statue'. That precisification (or those precisifications) makes the following come out true: 'If there were persisting statues, then the atoms arranged statuewise in the possession of the royal family long ago would compose the same statue as the atoms arranged statuewise owned by the folk ontologist

[82] It is also due, in part, to clauses (i) and (ii). 'Arranged statuewise' was defined partly in terms of atoms' being such that, if statues existed, then they would compose a statue. The non-vacuous truth of the counterpossible conditional in question—if statues existed, then the atoms would compose a statue—is grounded, in part, in what the folk mean by 'statue' (Ch. 1, §I). Thus—relying on arguments in Ch. 2 (§II) —it is in part a matter of convention whether atoms at a given time are arranged statuewise.

just quoted'. And, of course, if that comes out true, so does the claim that those atoms 'are arranged same-statuewise'.

So the courts can make a difference in whether 'arranged same-statuewise' is truly and determinately predicated of atoms existing at various times. This implies that the courts can make a difference in whether a sentence affirming a claim of statue identity over time is nearly as good as true. This, in turn, implies that—given eliminativism—the courts make a difference in whether it is 'true for practical purposes' that statue identity holds. Thus eliminativism can account for an element of convention in the 'truth for practical purposes' of claims about statue identity over time.[83]

Now imagine that folk ontology is true. Imagine statues really exist. I think it would then require magic for the courts to make a difference, just by their deliberations, in whether a statue has definitely survived some transformation. After all, suppose that transformation took place miles from, and even years before, the deliberations. Those deliberations could not—without the aid of the preternatural—reach back in time and across space to change what happened, making a difference in whether (or to what degree) some material object survived.

Suppose this is right. Suppose that folk ontology cannot (non-magically) accommodate the conventionalism in our practices regarding the identity over time of the inanimate macrophysical objects that it countenances. Eliminativism, we have seen, can easily accommodate this. So, given our supposition,

[83] Given eliminativism, conventional 'identity over time' is probably best interpreted as solving a 'coordination problem': a practical problem with a number of solutions, all on a par, which solutions require coordinating the actions of many people (see Lewis 1969, ch. 1). So, for example, we might secure certain benefits by legislating that 'this is the car I purchased' is true (for practical purposes) and everyone acting accordingly. 'Legislation' is needed since there are—prior to legislation—a variety of equally good ways (and equally nearly as good as true ways, since the folk meaning of 'same car' is not entirely precise) we could judge car identity for practical purposes.

our practice of treating certain cases of identity as practically somewhat conventional fits better with eliminativism than with folk ontology; in this way, eliminativism does better with some of our practices than does folk ontology.

I asked us to suppose that folk ontology requires magic in order to have conventional 'truth for practical purposes' of the relevant claims of identity over time. But while all may be willing to suppose this for the sake of argument, not all will really believe it. I anticipate two objections to this supposition.

To understand the first, recall I said that, given eliminativism, claims about the identity over time of alleged inanimate macrophysical objects are 'true for practical purposes' just in case they are nearly as good as true. And I added that, given folk ontology, claims about the identity over time of such objects are 'true for practical purposes' just in case they are plain old true. But one might object that I have unfairly limited the folk ontologist's options. One might say that neither the eliminativist *nor the folk ontologist* must make the literal truth of, for example, 'this is the same statue as the one I purchased' play the role of 'truth for practical purposes'. The folk ontologist should say that something else plays that role, this objection continues, something which could turn—in a non-magical way—on our conventions.

Much of the appeal of folk ontology is that it makes claims about the existence and persistence of inanimate macrophysical objects 'true for practical purposes' just because they are indeed true. It is in some sense self-defeating for the folk ontologist to try to divorce 'truth for practical purposes' from literal truth. And even if the folk ontologist insists on the divorce, it's not clear that it can be consummated. For there isn't an obvious replacement for literal truth out there.

It might *seem* like there is an obvious replacement. One might reply that, for the folk ontologist, the truth of whether *there are things arranged same-statuewise* could play the role of 'truth for practical purposes' of claims of statue identity. This

reply overlooks a crucial fact. If folk ontology is true, then whether atoms are arranged same-statuewise is a matter of whether they compose a single persisting statue. (Recall the account of 'same-statuewise' above.) Thus, if (and only if) folk ontology is true, whether atoms are arranged same-statuewise turns on convention only if whether a statue persists—and so whether a claim of statue identity is literally true—turns on convention. But it was just this result—the persistence of an actual statue turning on convention—that this reply was hoping to avoid. So this reply gets us nowhere.

There is a second objection to the claim that folk ontology renders spooky the conventional nature, for practical purposes, of the identity of inanimate macrophysical objects. This objection relies on an ontology of many co-located material objects with differing persistence conditions. A defender of this prodigal ontology could interpret the court's making determinate an otherwise borderline case of identity over time in the following manner:

> Suppose it is vague whether 'S at *t* is the same statue as S* at *t**' is true. This means that some, but not all, of the equally good candidates for being referred to by 'S at *t*'—which candidates are all co-located at *t*—are equally good candidates for being referred to by 'S* at *t**'. When the courts make 'S at *t* is identical with S* at *t**' determinately true, they do so not by magically reaching out and changing S or S*. They instead precisify 'S at *t*' so that all the candidates for being referred to by it are also candidates for being referred to by 'S* at *t**'. And in this way, whether it is 'true for practical purposes' that S at *t* is identical with S* at *t** can be a matter of convention.

Let us be clear about something. Full-blown conventionalism about persistence would imply that, for some object, whether that object persists is a matter of convention. The above account does not get us that (see Merricks, forthcoming

b). Rather, the above account says that there are many candidates—all of which presumably persist non-conventionally—and then makes some claims about the way our conventions secure the truth of certain sentences.

This is not an objection. Indeed, my own accommodation of conventional 'truth for practical purposes' for claims of persistence does not imply that there is some object such that it persists conventionally. Rather, I make claims about how the truth of sentences using predicates like 'arranged statuewise' turn on our conventions, and then parlay those claims into a thesis about how claims about persistence are 'true for practical purposes' as a matter of convention.

So the above account of 'conventional identity' and my account are alike in the following respects. Both suggest a particular way that the truth of some sentences can turn on our conventions. And both add that the truth of those sentences turning on our conventions in that way is good enough to make the 'truth for practical purposes' of certain claims of identity a matter of convention.

Despite their similarities, I prefer my account to that above. The above account postulates extensive co-location at a time, including—because this view is part of a wholly general approach that applies to all material objects—co-located persons. Moreover, for reasons made clear in Chapter 2 (§III), co-locationists ought to embrace perdurance and inconstant modal predicates. But I reject these concomitants of colocation as well as co-location itself.

My complaints about co-location (and its implications) are old news. But I have another objection to the account currently under discussion. This objection begins by considering a passage from Bernard Williams's 'The Self and the Future':

> [Conventionalist talk about identity over time] is the sort of thing indeed appropriate to lawyers deciding the ownership of some property which has undergone some bewildering

> set of transformations; they just have to decide, and in each situation, let us suppose, it has got to go to somebody, on as reasonable grounds as the facts and the law admit. But as a line to deal with a person's fears or expectations about his own future, it seems to have no sense at all. (1973: 61)

Pedants will object that Williams doesn't draw a perfectly clear distinction in this passage. For even the most avowed enemy of conventionalism about personal identity should grant that—as with property that has undergone bewildering transformations—lawyers could be forced to make a decision about personal identity. Suppose two infants in the hospital nursery fuse (providing the tabloids a nice contrast to their 'separated at birth' coverage). Two sets of parents then claim the fusion product. The courts must decide whose claims win out and so, let us suppose, who the fission product is. Surely the best procedure is to do so on as reasonable grounds as the facts and the law admit.

But I don't think Williams's point is that such an event could not happen. His point is, rather, that even if it did happen, the relevant facts of personal identity would still be nonconventional; the facts would be neither fixed nor determined nor constituted by the lawyers' decision or judicial decree. (So, for example, the courts could be mistaken.) Yet, Williams seems to claim, this is not so with respect to mere 'property'. The identity of such property can be somewhat conventional in that it can be at least partly constituted by a legal verdict.

I concur with this much of what Williams seems to be saying. As far as practical issues are concerned, the identity over time of things like statues can be somewhat a matter of convention. But claims about personal identity are never conventional, not even only in so far as practical matters are concerned. (Whether the person to be tortured tomorrow is identical with *me*—a matter of practical concern if ever there were one!—cannot be a matter of convention.)

We have here an intuitively compelling asymmetry. The 'truth for practical purposes' of claims of statue identity over time can be conventional; not so for the 'truth for practical purposes' of claims about personal identity over time. Any thesis that can justify this asymmetry thereby has a mark in its favour; any thesis that denies this asymmetry thereby has a mark against it.

The ontology of multiply co-located entities that we considered above can explain how claims of statue identity can be conventionally 'true for practical purposes'. But in doing so, it commits itself to an ontology according to which claims of personal identity over time can be 'true for practical purposes' as a matter of convention in just the same sense, and for just the same reasons, as can claims of statue identity over time.

By embracing both persons and statues, the co-locationist ontology implies that the identity over time of each is equally contingent on our conventions. And so, I believe, it is for folk ontology generally. Of course, some species of folk ontology will deny that the 'truth for practical purposes' of claims of personal identity can be conventional in any way at all. But I suspect that every species must say that such identity can be conventional for practical purposes in just the same way and to just the same extent as can statue identity. And so I suspect that every species of folk ontology denies the crucial asymmetry.

I can easily accommodate the asymmetry. The 'truth for practical purposes' of claims of identity over time for statues is, so I say, a matter of whether there are things arranged same-statuewise, which is to some extent conventional. But the 'truth for practical purposes' of personal identity over time is, I say, just their literal truth. That cannot be conventional, lest it be magical.

My ontology can explain and support the intuitive asymmetry between persons and statues regarding their identity over time being, for practical purposes at least, somewhat

conventional.[84] Folk ontology cannot. And so eliminativism not only accommodates our practices as well as folk ontology, but does so better than folk ontology.

One might object that our attitudes about conventional identity do not support my ontology generally. Suppose it turned out that, for example, atoms, molecules, and plants existed (perhaps they all have non-redundant causal powers). It could still be that we are intuitively inclined to think at least some of these enjoy, for practical purposes at least, conventional identity over time. Conversely, we might not think that the identity of *every* alleged object that I eliminate can be, for practical purposes, a matter of convention. Thus, the objection concludes, our native attitudes about conventional identity may not carve up the world in a way that mirrors my overall ontology.

I shall not dispute the substance of this objection. Rather, I deny its relevance to the point at hand. What is relevant—and what is untouched by this objection—is the following, which constitutes an advantage of my ontology over that of the folk. My ontology accommodates the clearest cases (those involving artefacts) of identity's being, for practical purposes, conventional. And my ontology accommodates the claim that personal identity fails to be conventional, even for practical purposes.

Consider puzzles regarding identity over time, puzzles like that involved in the Sorites Game or the Ship of Theseus. Consider also puzzles about essential properties, such as whether Kripke's (1980: 113) table could have been made of ice,

[84] I think *only* ontologies appropriately like mine can do this. Some might object that another way to do this is to start with Butler's (1736) 'loose and popular' identity for artefacts and 'strict and philosophical identity' for persons, endorsed also by Chisholm (1976, ch. 3). But I interpret Butler and Chisholm as eliminating artefacts but not persons. For they hold that there are persons (albeit simple ones), yet nothing that can gain or lose parts—and so nothing that is intuitively at all like a persisting artefact. (Cf. Chisholm's 1976: 97 ff. queries into whether familiar things like 'ships and trees and houses' are 'logical constructions'.)

or of a different block of wood. When considering such puzzles, I find two responses attractive. The first response, of the sort endorsed at the close of Chapter 2, is that these puzzles rest on a mistake. The puzzle-producing objects do not exist, so the puzzles disappear. The second response is that there are better and worse answers to these puzzles, but whether an answer is the best answer is somewhat conventional; once we have all the other facts, the best response to the puzzle is, in some sense, whatever we make it out to be.

These two responses are, if taken at face value, inconsistent. But we can interpret the second in such a way that it is consistent with the first, with the response defended at the end of Chapter 2. For example, I would say that, even though no ship is Theseus's (because there are no ships), we ought to decide which ship (if any) to treat as Theseus's ship as far as the practical matter of ownership is concerned. We ought to decide whether 'this ship belongs to Theseus' is 'true for practical purposes'. Moreover, I would add, our decision here—if it is both reasonable and we are apprised of all the other relevant facts—somehow constitutes what the 'truth for practical purposes' is. We cannot get it wrong.

All this seems plausible with respect to ships and, indeed, artefacts generally. But the conventionality of 'truth for practical purposes' is not plausible when it comes to matters of personal identity. This asymmetry, I argued, evidenced an advantage, with respect to practice, of my ontology over that of the folk. But there is a second way in which this asymmetry speaks in favour of my ontology. For—setting aside all questions of which ontology better accommodates our practices—I think this asymmetry indicates a fundamental difference between humans and alleged artefacts, such as statues. This fundamental difference, easily accounted for by my ontology, supports my contention that it is neither arbitrary nor unmotivated to believe in humans while eliminating statues.

IV. And Yet I Often Say 'There Are Statues'

I recently remarked to my 5-year-old daughter, while at a museum, 'there is a statue of a Roman emperor'. Was I lying? Did I say something false? No and no. For, I shall argue in this section, I—unlike the folk—do not speak falsely when saying things like 'there are statues' in the ordinary business of life.

When I say 'there is a statue of a Roman emperor', I mean that there are things arranged statue-of-a-Roman-emperorwise. Generally, when I say 'there is an F', when alleged Fs are supposed to be non-living macroscopica, I mean that there are things arranged F-wise. In such contexts, I am using 'there is' in a misleading or loose or even wrong way. I am using 'there is' deviantly. We can illustrate the idea here by replacing A, in one of our imagined exchanges above, with me. Here is how it would go:

TM. There are chairs over there.B. Do you mean to say that there are some things over there such that they are chairs? Do you affirm 'chairs exist' in the most literal and straightforward fashion possible?TM. I most certainly do not. You ask insightful questions.

The folk and I have different ontological beliefs. As a result, we often use the same sentences to express different propositions. But the proposition I express by saying 'there is a chair' in the ordinary business of life is closely related to the one the folk express by that same sentence. The former proposition is true if and only if the latter is nearly as good as true.

Suppose that you travel to the land of the Absolute Ptolemaists. (Recall that Absolute Ptolemaists believe the earth is absolutely fixed and the sun moves around it.) Although they speak English, you express a different proposition with 'the sun moved behind the elms' than do they. You say something true; they say something false. But when they

say 'meet us at the lagoon when the sun moves behind the elms', you have no trouble meeting them at the appointed time. The way you differ from them poses no practical difficulties whatsoever. Neither does the way the eliminativist differs from the folk.

In ordinary contexts when I say 'there is no chair in the room', I mean that there are no things arranged chairwise in the room. However, in contexts where ontology is at issue, I mean by 'there is no chair in the room' that there is no chair in the room. The phenomenon of a single speaker using the same words to mean different things in different context is, of course, very familiar. Suppose, while travelling in our imagined land, you stumble upon a fellow Heliocentrist. While marvelling with her that those around you believe the sun moves, you say 'the sun does not move behind the elms'. Yet this claim is consistent with what you meant earlier when you said to the Absolute Ptolemaists 'I was at the lagoon when the sun moved behind the elms'.

When I say 'there is a statue of a Roman emperor', I am using 'there is' deviantly and in such a way that what I say is true. Thus I assuage my conscience with respect to lying. But one might object that I am not so easily let off the hook. After all, I can be fairly certain that when I say 'there is a statue of a Roman emperor', most will interpret me folk-ontologically. I can be fairly certain that—given the truth of eliminativism—when I say such things I promote false beliefs in others. One might charge that I am like an adulterer who decides to use 'having an affair' to mean flying to the moon. He does not thereby avoid lying—or at least not its moral equivalent—when he says 'I am not having an affair'. (Compare: 'It all depends on what the meaning of "is" is'.)

To answer this charge, I return to the land of the Absolute Ptolemaists. There it seems that—unless you want to enter into a long discussion about astronomy—the *least misleading* thing you can say in ordinary contexts is something like 'when

the sun moves behind the elms, I'll meet you at the lagoon'. To say 'the sun will never move behind the elms' or 'if the sun moves behind the elms, then Russell is Pope' is to breed confusion. Likewise—unless I want to enter into a long discussion about ontology—the least misleading thing to say will often be 'there is a statue over there', even though I know my statement is likely to be misinterpreted as affirming what is only nearly as good as true. (All of this is in stark contrast to our imagined adulterer, who speaks in a misleading way when he could easily make the truth quite clear.)

Moreover, suppose I decided to 'speak plainly' and always say 'there are things arranged statuewise' instead of 'there are statues'. My plain speech presupposes a grasp of, and requires some sort of use of, the folk concept of *statue*. This is because knowing that there are things arranged statuewise requires us to know how things would be if—*per impossibile*—there were statues. (Things are arranged statuewise only if, if there were statues, then those things would compose a statue.)

We cannot accurately describe the world with speech that is 'more plain' than this. Although some Superbeing might be able to describe the contents of a museum solely in terms of the spatiotemporal and causal relations among microscopica, such a path is not open to us. To describe the contents of a museum, we must think and speak in terms of things arranged statuewise.

In Chapter 1 (§I) I said we could not understand what 'arranged statuewise' means without having the concept of *statue*. The point here is stronger. It is that many of the facts we grasp and express with the locution 'arranged statuewise' cannot be grasped or described by us unless we have the concept of *statue*. There are some facts about the world—such as the fact that there are things arranged statuewise—that we can neither express nor grasp without relying upon empty folk concepts.

Given that eliminativists must make use of empty folk concepts even in speaking plainly, we might as well, for convenience,

talk as if the concepts actually had application. At the very least, by speaking as if there are statues we are not forgoing a description of the truth that is *entirely free* of reliance on empty folk concepts. Such descriptions are not available to us.

Some might object that our needing concepts like that of *statue* in order to truly describe the world indicates that statues exist. But consider this: everything in my house is a nonunicorn. I understand this claim. I believe this claim. This claim is true. But I couldn't grasp it unless—given the obvious definition of 'non-unicorn'—I had the empty concept of *unicorn*. Those who think the indispensability of *statue* supports folk ontology are (or should be) on their way to believing in unicorns.

Moreover, consider a disposition like solubility. Suppose that some stuff is soluble if and only if were it submerged in water, then (*ceteris paribus*) it would dissolve. Suppose also that this disposition must be grounded in actual, categorical, non-subjunctive facts about that stuff. Finally, suppose we do not know what these facts are. In such a case, the only way we can describe certain truths about the world—such as that the stuff in question is soluble—is to invoke counterfactuals whose grounding is unknown to us.

Likewise, suppose that atoms are arranged statuewise only if, if there were statues, then those atoms would compose one. Suppose also that this counterpossible must be grounded in actual, categorical facts. (These facts would include facts about those atoms (including relational facts) and facts about what the folk mean by 'statue'.) Finally, suppose we cannot specify what these facts are. In such a case, the only way to report certain truths about the world—such as that there are atoms arranged statuewise—is to invoke counterpossibles whose grounding is unknown to us. I think that this is the way things are.

V. Conclusion

Eliminativism is true. And when the folk say 'there are statues', they ordinarily mean that there are statues. Thus the folk often say, and often believe, *falsehoods*. But false folk beliefs are *nearly as good as true*. Their being nearly as good as true makes them better, with respect to a number of epistemic norms, than beliefs like 'there are unicorns'.

Moreover, nearly as good as true folk beliefs are practically as good as true and, sometimes, even practically better than true. For eliminativism does better than standard folk ontology at accommodating our practice of treating certain cases of identity as somewhat conventional. Relatedly, my ontology does better than that of the folk at making sense of the intuitive asymmetry between artefacts and persons with respect to whether identity can be, for practical purposes, a matter of convention.

Folk concepts such as that of *statue,* although empty, are indispensable. That is, there are important truths about the world—practically important to us—that we could not grasp without them. Perhaps we could make do without grasping those truths. Perhaps we could make do without thinking in terms of statues or even in terms of things arranged statuewise. Perhaps we could abandon the folk-ontological framework, along with any other framework that is parasitic upon it, altogether. But to do so would be to abandon our way of life.

Eliminativism is true. But false folk-ontological beliefs are commendable in a variety of ways. And—even though eliminativism is true—empty folk-ontological concepts are indispensable, given our actual practical concerns. All of this speaks in favour of the ontology defended in these atoms arranged bookwise.

References

Alexander, Samuel (1920), *Space, Time, and Deity,* 2 vols. (London: Macmillan).

Alston, William P. (1990), 'Externalist Theories of Perception', *Philosophy and Phenomenological Research,* 50: 73–97.

Armstrong, D. M. (1997), *A World of States of Affairs* (Cambridge: Cambridge University Press).

Baker, Lynne Rudder (1997), 'Why Constitution is not Identity', *Journal of Philosophy,* 94: 599–621.

—— (2000), *Persons and Bodies: A Constitution View* (Cambridge: Cambridge University Press).

Barnes, Jonathan (ed.) (1995), *The Cambridge Companion to Aristotle* (Cambridge: Cambridge University Press).

Baxter, Donald L. M. (1988), 'Identity in the Loose and Popular Sense', *Mind,* 97: 575–82.

Benardete, José A. (1989), *Metaphysics: The Logical Approach* (New York: Oxford University Press).

Broad, C. D. (1925), *The Mind and its Place in Nature* (London: Routledge \& Kegan Paul).

Burke, Michael (1992), 'Copper Statues and Pieces of Copper', *Analysis,* 52: 12–17.

—— (1994), 'Dion and Theon: An Essentialist Solution to an Ancient Puzzle', *Journal of Philosophy,* 91: 129–39.

Butler, Joseph (1736), 'Of Personal Identity', first appendix of *The Analogy of Religion*; repr. in John Perry (ed.) *Personal Identity* (Berkeley: University of California Press, 1975).

Carter, W. R. (1988), 'Our Bodies, Our Selves', *Australasian Journal of Philosophy,* 66: 308–19.

—— (1999), 'Will I Be a Dead Person?', *Philosophy and Phenomenological Research,* 59: 167–71.

Cartwright, Richard (1975), 'Scattered Objects', in Keith Lehrer (ed.), *Analysis and Metaphysics* (Dordrecht: Reidel).

Chalmers, David (1996), *The Conscious Mind* (New York: Oxford University Press).
Chisholm, Roderick M. (1976), *Person and Object* (LaSalle, Ill.: Open Court).
Cover, Jan and O'Leary-Hawthorne, John (1996), 'Free Agency and Materialism', in Jeff Jordan and Daniel Howard-Snyder (eds.), *Faith, Freedom, and Rationality* (Rowman \& Littlefield).
Crisp, Thomas M. and Warfield, Ted A. (2000), 'The Irrelevance of Indeterministic Counterexamples to Principle Beta', *Philosophy and Phenomenological Research,* 61: 173–84.
Doepke, Frederick (1982), 'Spatially Coinciding Objects', *Ratio,* 24: 45–60; repr. in Rea (1997*b*).
Dummett, Michael (1975), 'Wang's Paradox', *Synthese,* 25: 301–24.
Elder, Crawford (1996), 'On the Reality of Medium-Sized Objects', *Philosophical Studies,* 83: 191–211.
Evans, Gareth (1978), 'Can there be Vague Objects?', *Analysis,* 38: 208.
Finch, Alicia and Warfield, Ted A. (1998), 'The *Mind* Argument and Libertarianism', *Mind* 107: 515–28.
Fodor, Jerry (1989), 'Making Mind Matter More', *Philosophical Topics,* 17: 59–79.
Francescotti, Robert (1999), 'How to Define Intrinsic Properties', *Noûs,* 33: 590–609.
Geach, Peter (1980), *Reference and Generality,* 3rd edn. (Ithaca, NY: Cornell University Press).
Gibbard, Allan (1975), 'Contingent Identity', *Journal of Philosophical Logic,* 4: 187–221; repr. in Rea (1997*b*).
Ginet, Carl (1966), 'Might We Have No Choice?', in Keith Lehrer (ed.), *Freedom and Determinism* (New York: Random House).
Harman, Gilbert (1973), *Thought* (Princeton: Princeton University Press).
Hawley, Katherine (1998), 'Merricks on whether being Conscious is Intrinsic', *Mind,* 107: 841–3.
Heller, Mark (1990), *The Ontology of Physical Objects* (Cambridge: Cambridge University Press).
—— (1996), 'Against Metaphysical Vagueness', in J. Tomberlin (ed.), *Philosophical Perspectives,* 10 (Oxford: Blackwell).
—— (2000), 'Temporal Overlap is not Coincidence', *Monist,* 83: 362–80.

Hestevold, H. Scott (1980–1), 'Conjoining', *Philosophy and Phenomenological Research,* 41: 371–85.
Hirsch, Eli (1993), 'Peter van Inwagen's *Material Beings*', *Philosophy and Phenomenological Research,* 53: 687–91.
Honderich, Ted (1988), *A Theory of Determinism* (Oxford: Oxford University Press).
Horgan, Terence (1985), 'Compatibilism and the Consequence Argument', *Philosophical Studies,* 47: 339–56.
Johnston, Mark (1992), 'Constitution is not Identity', *Mind,* 101: 89–105.
Kane, Robert (1998), *The Significance of Free Will* (New York: Oxford University Press).
Kim, Jaegwon (1982), 'Psychophysical Supervenience', *Philosophical Studies,* 41: 51–70; repr. in Kim (1993*b*).
—— (1989*a*), 'Mechanism, Purpose, and Explanatory Exclusion', in J. Tomberlin (ed.) *Philosophical Perspectives,* 3 (Atascadero, Calif.: Ridgeview); repr. in Kim (1993*b*).
—— (1989*b*), 'The Myth of Nonreductive Materialism', *Proceedings and Addresses of the American Philosophical Association,* 63: 31–47.
—— (1993*a*), 'The Nonreductivist's Troubles with Mental Causation', in J. Heil and A. Mele (eds.), *Mental Causation* (Oxford: Oxford University Press); repr. in Kim (1993*b*).
—— (1993*b*), *Supervenience and Mind* (Cambridge: Cambridge University Press).
—— (1998) *Mind in a Physical World: An Essay on the Mind–Body Problem and Mental Causation* (Cambridge, Mass.: MIT Press).
Kornblith, Hilary (1993), *Inductive Inference and its Natural Ground: An Essay in Naturalistic Epistemology* (Cambridge, Mass.: MIT Press).
Kripke, Saul (1980), *Naming and Necessity* (Cambridge, Mass.: Harvard University Press).
Langton, Rae and Lewis, David (1998), 'Defining "Intrinsic" ', *Philosophy and Phenomenological Research,* 58: 333–45.
Lehrer, Keith and Paxson, Thomas (1969), 'Knowledge: Undefeated Justified True Belief', *Journal of Philosophy,* 66: 225–37.
Leonard, Henry S. and Goodman, Nelson (1940), 'The Calculus of Individuals and its Uses', *Journal of Symbolic Logic,* 5: 45–55.
Levison, Arnold (1978), 'Chisholm and "the Metaphysical Problem of Human Freedom" ', *Philosophia,* 8: 537–41.

Lewis, David (1968), 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy,* 65: 113–26; repr, with postscripts, in Lewis (1983).
—— (1969), *Convention: A Philosophical Study* (Cambridge, Mass.: Harvard University Press).
—— (1971), 'Counterparts of Persons and their Bodies', *Journal of Philosophy,* 68: 203–11; repr. in Lewis (1983).
—— (1973*a*), 'Counterfactuals and Comparative Possibility', *Journal of Philosophical Logic,* 2: 418–46; repr. in Lewis (1986*b*).
—— (1973*b*), *Counterfactuals* (Oxford: Blackwell).
—— (1976), 'Survival and Identity', in Amelie Rorty (ed.), *The Identities of Persons* (Berkeley: University of California Press); repr., with postscripts, in Lewis (1983).
—— (1978), 'Truth in Fiction', *American Philosophical Quarterly,* 15: 37–46; repr., with postscripts, in Lewis (1983).
—— (1981), 'Are we Free to Break the Laws?', *Theoria,* 47: 113–21. repr. in Lewis (1986*b*).
—— (1983), *Philosophical Papers, vol. I* (New York: Oxford University Press).
—— (1986*a*), *On the Plurality of Worlds* (Oxford: Blackwell).
—— (1986*b*), *Philosophical Papers, vol. II* (New York: Oxford University Press).
—— (1991), *Parts of Classes* (Oxford: Blackwell).
—— (1993), 'Many, but Almost One', in Keith Campbell, John Bacon, and Lloyd Reinhardt (eds.), *Ontology, Causality, and Mind: Essays on the Philosophy of D. M. Armstrong* (Cambridge: Cambridge University Press); repr. in Lewis (1999).
—— (1999), *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press).
Locke, John (1975), *An Essay concerning Human Understanding,* ed. P. H. Nidditch (Oxford: Clarendon Press).
Lowe, E. J. (1983), 'Instantiation, Identity, and Constitution', *Philosophical Studies,* 44: 45–59.
—— (1989), *Kinds of Being* (Oxford: Blackwell).
—— (1996), *Subjects of Experience* (Cambridge: Cambridge University Press).
Lycan, William (1979), 'The Trouble with Possible Worlds', in Michael Loux (ed.), *The Possible and the Actual* (Ithaca, NY: Cornell University Press).

McKay, Thomas and Johnson, David (1996), 'A Reconsideration of an Argument against Compatibilism', *Philosophical Topics,* 24: 113–22.

Mackie, Penelope (1993), 'Ordinary Language and Metaphysical Commitment', *Analysis,* 53: 243–51.

McLaughlin, Brian (1989), 'Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical', in J. Tomberlin (ed.), *Philosophical Perspectives,* 3 (Atascadero, Calif.: Ridgeview).

Malcolm, Norman (1968), 'The Conceivability of Mechanism', *Philosophical Review,* 77: 45–72.

Markosian, Ned (1998), 'Brutal Composition', *Philosophical Studies,* 92: 211–49.

Maudlin, Tim (1994), *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics* (Oxford: Blackwell).

Merricks, Trenton (1994), 'Endurance and Indiscernibility', *Journal of Philosophy,* 91: 165-184.

—— (1995*a*), 'On the Incompatibility of Enduring and Perduring Entities', *Mind,* 104: 523–31.

—— (1995*b*), 'Warrant Entails Truth', *Philosophy and Phenomenological Research,* 55: 851–55.

—— (1997*a*), 'Fission and Personal Identity over Time', *Philosophical Studies,* 88: 163–86.

—— (1997*b*), 'More on Warrant's Entailing Truth', *Philosophy and Phenomenological Research,* 57: 627–31.

—— (1998*a*), 'Against the Doctrine of Microphysical Supervenience', *Mind,* 107: 59–71.

—— (1998*b*), 'There are no Criteria of Identity over Time', *Noûs,* 32: 106–24.

—— (1998*c*), 'On whether being Conscious is Intrinsic', *Mind,* 107: 845–6.

—— (1999*a*), 'Persistence, Parts, and Presentism', *Noûs,* 33: 421–38.

—— (1999*b*), 'Composition as Identity, Mereological Essentialism, and Counterpart Theory', *Australasian Journal of Philosophy,* 77: 192–5.

—— (1999*c*), 'Endurance, Psychological Continuity, and the Importance of Personal Identity', *Philosophy and Phenomenological Research,* 54: 983–97.

—— (2000*a*) ' "No Statues" ', *Australasian Journal of Philosophy,* 78: 47–52.

Merricks, Trenton (2000*b*), 'Perdurance and Psychological Continuity', *Philosophy and Phenomenological Research,* 61: 195–8.

—— (2001), 'Varieties of Vagueness', *Philosophy and Phenomenological Research,* 62: 145–57.

—— (forthcoming *a*), 'Physicalism and Immortality: How to Live Forever without Saving your Soul' in Kevin Corcoran (ed.), *Persons, Bodies, and the Prospect of Survival: Essays on the Metaphysics of Human Persons* (Ithaca, NY: Cornell University Press).

—— (forthcoming *b*), 'Realism about Personal Identity over Time', in J. Tomberlin (ed.), *Philosophical Perspectives,* 15 (Oxford: Blackwell).

Michael, Michaelis and O'Leary-Hawthorne, John (1996), 'Compatibilist Semantics in Metaphysics: A Case Study', *Australasian Journal of Philosophy,* 74: 117–34.

Mill, J. S. (1843), *A System of Logic* (London: Longmans, Green, Reader, and Dyer).

Mills, Eugene (1996), 'Interactionism and Overdetermination', *American Philosophical Quarterly,* 33: 105–17.

Morgan, C. Lloyd (1923), *Emergent Evolution* (London: Williams \& Norgate).

Nagel, Thomas (1986), *The View from Nowhere* (New York: Oxford University Press).

Noonan, Harold (1989), *Personal Identity* (London: Routledge).

—— (1999), 'Identity, Constitution, and Microphysical Supervenience', *Proceedings of the Aristotelian Society,* 99: 273–80.

O'Connor, Timothy (1995), 'Agent Causation', in Timothy O'Connor (ed.), *Agents, Causes, and Events: Essays on Indeterminism and Free Will* (New York: Oxford University Press).

—— (2000), *Persons and Causes: The Metaphysics of Free Will* (Oxford: Oxford University Press).

Olson, Eric T. (1995), 'Why I have no hands', *Theoria,* 61: 182–97.

—— (1997), *The Human Animal: Personal Identity without Psychology* (New York: Oxford University Press).

Parfit, Derek (1984), *Reasons and Persons* (Oxford: Clarendon Press).

Plantinga, Alvin (1974), *The Nature of Necessity* (Oxford: Clarendon Press).

—— (1993), *Warrant and Proper Function* (Oxford: Oxford University Press).

Pollock, John (1974), *Knowledge and Justification* (Princeton: Princeton University Press).
—— (1986), *Contemporary Theories of Knowledge* (Totowa, NJ: Rowman \& Littlefield).
Quine, W. V. (1953), 'On What There Is', in *From a Logical Point of View* (Cambridge, Mass.: Harvard University Press).
Rea, Michael C. (1995), 'The Problem of Material Constitution', *Philosophical Review,* 104: 525–52.
—— (1997*a*), 'Supervenience and Co-location', *American Philosophical Quarterly,* 34: 367–75.
—— (ed.) (1997*b*), *Material Constitution: A Reader* (New York: Rowman \& Littlefield).
—— (1998*a*), 'In Defense of Mereological Universalism', *Philosophy and Phenomenological Research,* 58: 347–60.
—— (1998*b*), 'Sameness without Identity: An Aristotelian Solution to the Problem of Material Constitution', *Ratio,* 11: 316–28; repr. in David S. Oderberg (ed.) *Form and Matter: Contemporary Themes in Metaphysics* (Oxford: Blackwell, 1999).
—— (2000), 'Constitution and Kind Membership', *Philosophical Studies,* 97: 169–93.
Richards, Tom (1975), 'The Worlds of David Lewis', *Australasian Journal of Philosophy,* 53: 105–18.
Rosen, Gideon (1990), 'Modal Fictionalism', *Mind,* 99: 327–54.
Rosenberg, Jay (1993), 'Comments on Peter van Inwagen's *Material Beings*', *Philosophy and Phenomenological Research,* 53: 701–8.
Russell, Bertrand (1923), 'Vagueness', *Australasian Journal of Philosophy and Psychology,* 1: 84–92.
Ryle, Gilbert (1949), *The Concept of Mind* (Chicago: University of Chicago Press).
Searle, John (1992), *The Rediscovery of the Mind* (Cambridge, Mass.: MIT Press).
Shoemaker, Sydney (1997), 'Parfit on Identity', in Jonathan Dancy (ed.), *Reading Parfit* (Oxford: Blackwell).
—— (1999*a*), 'Critical Notice of Eric Olson's *The Human Animal*', *Noûs,* 33: 496–504.
—— (1999*b*), 'Self, Body, and Coincidence', *Proceedings of the Aristotelian Society,* suppl. vol. 73: 287–306.
Sidelle, Alan (1998), 'A Sweater Unraveled: Following One Thread of Thought for Avoiding Coincident Entities', *Noûs,* 32: 423–48.

Sider, Theodore (1997), 'Four-Dimensionalism', *Philosophical Review,* 106: 197–231.
—— (1999*a*), 'Presentism and Ontological Commitment', *Journal of Philosophy,* 96: 325–47.
—— (1999*b*), 'Merricks on Microphysical Supervenience', unpub.
Skyrms, Brian (1976), 'Possible Worlds, Physics, and Metaphysics', *Philosophical Studies,* 30: 323–32.
Sorensen, Roy (1988), *Blindspots* (Oxford: Clarendon Press).
Sosa, Ernest (1987), 'Subjects among Other Things', in J. Tomberlin (ed.), *Philosophical Perspectives,* 1 (Atascadero, Calif.: Ridgeview); repr. in Rea (1997*b*).
Strawson, P. F. (1959), *Individuals* (London: Methuen).
Swinburne, Richard (1986), *The Evolution of the Soul* (Oxford: Clarendon Press).
—— (1995), 'Thisness', *Australasian Journal of Philosophy,* 73: 389–400.
Teller, Paul (1989), 'Relativity, Relational Holism, and the Bell Inequalities', in J. Cushing and E. McMullin (eds.), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem* (Notre Dame, Ind.: University of Notre Dame Press).
Unger, Peter (1980), 'The Problem of the Many', in P. French, T. Uehling Jr., and H. Wettstein (eds), *Midwest Studies in Philosophy,* vol. v (Minneapolis: University of Minnesota Press).
Vallentyne, Peter (1997), 'Intrinsic Properties Defined', *Philosophical Studies,* 88: 209–19.
van Cleve, James (1986), 'Mereological Essentialism, Mereological Conjunctivism, and Identity through Time', in P. French, T. Vehling Jr., and H. Wettstein (eds.), *Midwest Studies in Philosophy,* vol. xi (Minneapolis: University of Minnesota Press).
van Inwagen, Peter (1981), 'The Doctrine of Arbitrary Undetached Parts', *Pacific Philosophical Quarterly,* 62: 123–37; repr. in Rea (1997*b*).
—— (1983), *An Essay on Free Will* (Oxford: Clarendon Press).
—— (1990), *Material Beings* (Ithaca, NY: Cornell University Press).
—— (1993), 'Reply to Reviewers [of *Material Beings*]', *Philosophy and Phenomenological Research,* 53: 709–19.
—— (1998), 'Meta-Ontology', *Erkenntnis,* 48: 233–50.
Wheeler, Samuel (1986), 'Persons and their Micro-Particles', *Noûs,* 20: 333–49.
Widerker, David (1987), 'On an Argument for Incompatibilism', *Analysis,* 47: 37–41.

Wiggins, David (1968), 'On Being in the Same Place at the Same Time', *Philosophical Review,* 77: 90–5; repr. in Rea (1997*b*).

—— (1973), 'Towards a Reasonable Libertarianism', in T. Honderich (ed.), *Essays on Freedom of Action* (London: Routledge \& Kegan Paul).

Williams, Bernard (1973), *Problems of the Self* (Cambridge: Cambridge University Press).

Williamson, Timothy (1994), *Vagueness* (London: Routledge).

Wittgenstein, Ludwig (1961), *Tractatus Logico-Philosophicus,* trans. D. F. Pears and B. F. McGuinness (London: Routledge \& Kegan Paul).

Yablo, Stephen (1992), 'Mental Causation', *Philosophical Review,* 101: 245–80.

—— (1997), 'Wide Causation', in J. Tomberlin (ed.), *Philosophical Perspectives,* 11 (Oxford: Blackwell).

Zimmerman, Dean (1995), 'Theories of Masses and Problems of Constitution', *Philosophical Review,* 104: 53–110.

Index

Alexander, Samuel 65, 90 n.
Alston, William 65 n. 8
Aristotle 169 n. 4
Armstrong, D. M. 20
arranged same-statuewise, defined 176–7
arranged statuewise, defined 4
Augustine 47
Baker, Lynne Rudder 39 n. 3, 40, 40 n.
Barnes, Jonathan 169 n. 4
Baxter, Donald 20
Benardete, José 169
biological persistence conditions for persons 86–7, 133
Broad, C. D. 62, 90 n., 112 n., 139
Burke, Michael 38 n., 39 n. 4, 40 n., 101
Butler, Joseph 184 n.
Carter, W. R. 53, 86
Cartwright, Richard 8 n.
causal irrelevance, defined 57, 139 n. 1
Causal Principle 58
Chalmers, David 50, 143 n.
Chisholm, Roderick 65 n. 8, 184 n.
Cicero 158
co-location 33–47, 53, 77–8, 82–3, 85–6, 124, 130–4, 180–1, 183; defined 39
composition as identity 17 n., 20–8, 31, 56 n., 64
Consciousness (C) 94
constitution 20 n. 11, 39 n. 3
Copernicus 165 n. 3
counterpart theory 25–7, 46–7
counterpossibles 5–7, 36–7, 177, 189
Cover, Jan 158, 160 n.
Crisp, Thomas 156 n.
Descartes, René 47, 120, 135
Doepke, Frederick 39 n. 3
downward causation 60–1, 62 n. 5, 84, 109–10, 113, 148–54, 159–61
Dummett, Michael 33
Elder, Crawford 77 n. 11
emergent properties 62, 90 n., 111 n., 140
endurance 23, 27, 82 n.
Evans, Gareth 33
Exclusion Argument 139
'exists'; deviant uses of 18–19, 163–8; univocity of 19 n., 168–70
Finch, Alicia 156 n.
Fodor, Jerry 138
four-dimensionalism, *see* perdurance
Francescotti, Robert 92 n. 7
free will 84, 155–61
Geach, Peter 27, 38 n.
Gibbard, Allan 26 n., 44
Ginet, Carl 155
Goldman, Alvin 65 n. 8

Goodman, Nelson 8 n.
Grice, H. P. 65 n. 8
Guinness, draught 63
Harman, Gilbert 75 n.
Hawley, Katherine 101 n. 11
Hawthorne, John 158, 160 n., 165 n. 2–3
Heller, Mark 32, 34, 38 n., 50
Hestevold, H. Scott 76 n. 11
Hirsch, Eli 13
Honderich, Ted 160 n.
Horgan, Terence 156
identity; contingent 24–7; conventional 176–85, 190; over time, *see* endurance; perdurance; relative 26–7, 38; vague 33
Johnson, David 156 n.
Johnston, Mark 40 n.
Jordan, Michael 108, 147
Kane, Robert 156 n., 158, 160 n.
Kim, Jaegwon 57 n. 2, 64, 65, 65 n. 7, 71, 92 n. 6, 139, 143
Kornblith, Hilary 11–12
Kripke, Saul 27, 184
Langton, Rae 92 n. 6
Lehrer, Keith 75 n.
Leibniz, G. W. 6
Leonard, Henry S. 8 n.
Levison, Arnold 160 n.
Lewis, David 6, 8 n., 17 n., 20–1, 23 n. 15, 25, 26 n., 33, 36, 46, 76, 92 n. 6, 95, 98, 156, 168, 178 n.
Locke, John 40, 47–50
Lowe, E. J. 40 n., 122 n.
Lycan, William 76 n. 10
McKay, Thomas 156 n.
Mackie, Penelope 165 n. 2
McLaughlin, Brian 139 n. 2
Malcolm, Norman 71, 139
Markosian, Ned 76 n. 11
Maudlin, Tim 62 n. 6
mereological essentialism 24–5, 27, 38, 43, 53–4, 77 n. 11, 96
Michael, Michaelis 165 n. 2–3
Micro Exclusion Argument 141
Microphysical Closure 141
Mill, J. S. 90 n.
Mills, Eugene 66
modal predicates, inconstant 46–7, 77–8, 181
Moore, G. E. 12
Morgan, C. Lloyd 90 n.
Nagel, Thomas 135 n.
nearly as good as true, defined 171
Noonan, Harold 50, 101, 165 n. 2
O'Connor, Timothy 156 n., 160 n.
O'Leary-Hawthorne, John, *see* Hawthorne, John
Olson, Eric 49, 86
Overdetermination Argument 56
Parfit, Derek 35, 125 n., 135 n.
Paxson, Thomas 75 n.
perdurance 5, 21 n. 13, 22–3, 44–7, 48 n. 10, 53, 77–8, 97–9, 181
Plantinga, Alvin 27, 75 n.
Plato 47
Pollock, John 75 n.
Ptolemy 165 n. 3
Quine, W. V. 168
Rea, Michael 38 n., 41 n. 6–7, 77 n. 12, 86 n.
reductionism 11–12, 20, 28, 35
Richards, Tom 76 n. 10

Rosen, Gideon 6
Rosenberg, Jay 7
Russell, Bertrand 33, 188
Ryle, Gilbert 168–9
schema of the Overdetermination Argument 79–80, 112–13
Searle, John 11–12, 20
Ship of Theseus 5, 41 n. 7, 184–5
Shoemaker, Sydney 53, 86 n., 122 n.
Sidelle, Alan 38 n.
Sider, Theodore 5, 8 n., 101–3, 171 n.
Skyrms, Brian 76 n. 10
Sorensen, Roy 34
sorites paradox 32–8, 53–4, 124–30
Sosa, Ernest 39 n. 4, 53
Step One argument 89
Strawson, P. F. 122 n.
Swinburne, Richard 11–12, 48 n. 9
Teller, Paul 62 n. 6
temporal parts, *see* perdurance
Unger, Peter 95
unrestricted composition 8, 16–17, 51, 74–8
vagueness 33–7, 53–4, 124–30
Vallentyne, Peter 92 n. 6
van Cleve, James 38 n., 77 n. 12
van Inwagen, Peter 2 n., 15, 38 n., 41, 52, 53, 86, 155, 156 n., 160 n., 162–8, 171
Warfield, Ted A. 156 n.
Wheeler, Samuel 126
Widerker, David 156 n.
Wiggins, David 40, 155
Williams, Bernard 181–2
Williamson, Timothy 34
Wittgenstein, Ludwig 168
Yablo, Stephen 139, 143 n.
Zimmerman, Dean 31, 38 n., 39 n. 4, 127 n.